



Construction de courbes de fragilité sismique par la représentation de Karhunen-Loève

Fabien Giraudeau

► To cite this version:

Fabien Giraudeau. Construction de courbes de fragilité sismique par la représentation de Karhunen-Loève. Médecine humaine et pathologie. Université Blaise Pascal - Clermont-Ferrand II, 2015. Français. NNT : 2015CLF22540 . tel-01248250

HAL Id: tel-01248250

<https://theses.hal.science/tel-01248250>

Submitted on 4 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : D.U. 2540

EDSPIC : 684

UNIVERSITE BLAISE PASCAL - CLERMONT II

Ecole Doctorale

Sciences pour l'Ingénieur de Clermont-Ferrand

Thèse

Présentée par

Fabien Giraudeau

Ingénieur des Arts et Manufactures - Agrégé de Mathématiques

pour obtenir le grade de

Docteur d'Université

SPECIALITE : Matériaux, Structure, Fiabilité

Construction de courbes de fragilité sismique par la représentation de Karhunen-Loève

Soutenue publiquement le 8 janvier 2015 à Clamart (92) devant le jury :

Pr. Frédéric Ragueneau (ENS Cachan)

M. Christian Cremona (CEREMA)

M. Jean-Marc Martinez (CEA)

M. Cyril Feau (CEA)

Mme Irmela Zentner (EDF)

Pr. Michel Fogli (Institut Pascal)

Président

Rapporteur

Rapporteur

Invité

Invitée

Directeur de thèse

Mots-clés

Courbe de fragilité - Karhunen-Loève - Enrichissement - Tri par classes

Résumé

La probabilité de défaillance d'une structure sous séisme est représentée à l'aide de « courbes de fragilité ». Pour les estimer, nous proposons d'enrichir une base de données pré-existante à l'aide du modèle de l'article [53], qui se fonde sur la représentation de Karhunen-Loève. Les signaux générés sont triés par classes d'indicateur de nocivité sismique à l'aide d'un algorithme de partitionnement de données. Malgré la ressemblance certaine que présentent plusieurs signaux simulés, et les conséquences que nous tirons de ce problème, ils sont soumis à la structure. Les signaux de réponses résultants sont eux aussi enrichis, en prenant en compte certaines incertitudes afin de construire un intervalle autour de la courbe. La méthode fonctionne pour tout indicateur de nocivité sismique, et permet de s'affranchir de plusieurs hypothèses simplificatrices courantes. Les caractéristiques du scénario sismique sont conservées lors de l'enrichissement, et le processus modélisant le mouvement du sol garde toute sa généralité. Notre démarche est validée d'abord sur un cas simple, puis sur un cas industriel.

« Construction of seismic fragility curves with the Karhunen-Loève expansion »

Keywords

Fragility curves - Karhunen-Loève - Enrichment - Class sorting

Abstract

The failure probability of a structure under earthquake is represented with « fragility curves ». To estimate them, we propose to enrich a pre-existing data basis using the model of the article [53], which is based on the Karhunen-Loeve expansion. The generated signals are sorted by seismic indicator classes using a data partitioning algorithm. Despite the resemblance between some simulated signals, and the consequences we draw from this problem, the structure is submitted to them. The resulting response signals are also enriched, taking into account uncertainties to construct an interval around the curve. The method works for any seismic indicator, and overcomes several common simplifying assumptions. The characteristics of the seismic scenario are preserved during the enrichment, and the process modeling the ground motion retains its generality. Our approach is first validated on a simple case, then on an industrial case.

Laboratoire d'accueil :

LaMSID - UMR EDF/CNRS/CEA 8193

EDF R&D 1, avenue du Général de Gaulle 92141 Clamart Cedex

Tel : +33 (0)1 47 65 45 08

Fax : +33 (0)1 47 65 54 14

Directeur : Patrick Massin

Encadrants : Irmela Zentner (EDF) et Cyril Feau (CEA)

Remerciements

Ce travail, financé par le CNRS, EDF, le CEA et AREVA, n'aurait pu voir le jour sans l'expertise de mon directeur, le Professeur Michel Fogli ; pas plus qu'il n'aurait avancé correctement sans mes encadrants Irmela Zentner et Cyril Feau. Mes remerciements leur sont tout particulièrement destinés.

Les deux rapporteurs de ce travail, Christian Cremona et Jean-Marc Martinez, ont apporté leurs précieuses lumières à la clôture de ce travail ; à eux, ainsi qu'au président de mon jury de soutenance le Professeur Frédéric Ragueneau, j'exprime ma sincère gratitude. Je tiens également à remercier Catherine Berge-Thierry, Evelyne Foerster, Laurent Borsoi, Fabrice Poirion et Jean-Michel Thiry, de leur participation au long de ces années.

Parmi mes amis du LaMSID, j'ai une dette particulière envers son directeur Patrick Massin, ainsi qu'à Claude Stolz et Alexandre Martin, qui m'ont régulièrement éclairé de leur culture scientifique. Les autres sont si nombreux, la crainte d'en oublier si forte, que je les salue ici collectivement, tout en formulant une dédicace particulière à Marie-Line.

Je conclurai en dédiant de tout mon cœur ce travail à deux personnes qui se seraient réjouis d'assister à ma soutenance.

Table des matières

Introduction et démarche	9
1 Bibliographie	13
1.1 Caractérisation du mouvement du sol	13
1.2 Modèles stochastiques de mouvement sismique	15
1.2.1 Modèles classiques paramétriques	16
1.2.2 Modèle empirique reposant sur la représentation de Karhunen-Loève	19
1.3 Les courbes de fragilité	19
1.3.1 Définition et généralités	19
1.3.2 Le modèle log-normal	20
1.3.3 Le <i>scaling</i>	21
1.4 Conclusion	21
2 Méthodologie	23
2.1 Enrichissement d'une famille de signaux : « méthode K- L »	23
2.1.1 Une forme du théorème de Karhunen-Loève	24
2.1.2 Identification de l'excitation : « modèle K-L »	24
2.2 Estimation d'une densité conjointe	26
2.2.1 Développement sur une base de polynômes orthogonaux	27
2.2.2 L'estimateur à noyaux	29
2.3 Simulation de variables dépendantes	31
2.4 Tri par classes et estimation des courbes de fragilité	36
2.5 Prise en compte des incertitudes	38
2.5.1 Incertitude épistémique	38
2.5.2 Incertitude statistique	38
2.5.3 Incertitudes sur le modèle K-L	39
3 Analyse et validation de la méthodologie sur un cas simple	45
3.1 Données	45
3.1.1 Base de signaux réels	45

3.1.2	Modèles mécaniques	46
3.2	Représentativité pour la simulation de signaux d'entrée	48
3.2.1	Paramètres statistiques globaux	48
3.2.2	Modes propres de Karhunen-Loève	52
3.2.3	Taille de base intermédiaire	53
3.3	Tri par classes	56
3.3.1	Corrélation intra-classes des signaux d'entrée	56
3.3.2	Détermination de la taille des classes	58
3.4	Représentativité pour la simulation de signaux de réponse	59
3.4.1	Corrélation intra-classes des signaux de réponse	59
3.4.2	Paramètres statistiques globaux	60
3.4.3	Courbes de fragilité	61
3.5	Prise en compte de l'incertitude	65
3.5.1	Incertaince statistique	65
3.5.2	Incertaince sur le modèle K-L	67
3.6	Discussion	70
4	Application à un cas réel	73
4.1	Présentation de la ligne	73
4.2	Calcul de courbes de fragilité	79
4.3	Conclusion	84
A	Rappels de probabilités réelles	87
A.1	Variables aléatoires	87
A.2	Vecteurs aléatoires	88
A.3	Processus stochastiques	89
A.3.1	Généralités	89
A.3.2	Processus stationnaires	90
B	Echantillonnage	93
C	Démonstration du théorème de Karhunen-Loève	97

Introduction et démarche

Les Etudes Probabilistes de Sûreté (EPS) cherchent à estimer les risques liés aux installations industrielles. Elles considèrent des « événements initiateurs », et leurs fréquences d'occurrence annuelle, pour évaluer les risques en termes d'« événements redoutés ». Ceux-ci caractérisent la défaillance de l'installation et peuvent signifier, dans le cas de l'industrie nucléaire, fusion du cœur ou rejet de produits de fission hors de l'enceinte de confinement.

Si l'événement initiateur est le séisme, la mise en œuvre d'une EPS pour une installation et un site donnés comporte quatre étapes fondamentales [86] :

1. la construction de la courbe d'« aléa sismique », qui représente la probabilité de dépassement d'une valeur caractéristique de l'intensité sismique au cours d'une période de référence¹. L'intensité sismique est évaluée à l'aide d'un indicateur de nocivité² ;
2. l'établissement des arbres de défaillances, représentations graphiques des relations de cause à effet qui conduisent aux événements redoutés ;
3. la construction des **courbes de fragilité sismique** qui représentent, pour chaque structure de l'installation, sa probabilité de défaillance conditionnelle en fonction de l'intensité sismique à laquelle elle est soumise ;
4. le calcul des fréquences d'occurrence des événements redoutés.

Cette thèse s'intéresse au troisième point de cette démarche : l'estimation de la courbe de fragilité d'une structure sous séisme.

Cette courbe est croissante et varie de zéro à un, valeurs qui correspondent respectivement aux séismes de faibles et fortes intensités. Mais ceci n'impose pas *a priori* qu'elle soit monotone. Certaines caractéristiques de l'excitation, comme son contenu fréquentiel, peuvent en effet dépendre de son intensité et empêcher la monotonie de la courbe. Le comportement potentiellement non-linéaire de la structure est également un facteur de perte de monotonie de la courbe. Nous aurons l'occasion d'y revenir au cours de ce mémoire.

La courbe de fragilité sismique est censée intégrer l'ensemble des incertitudes inhérentes au problème traité : incertitudes liées au chargement sismique, aux comportements des structures porteuses dans le cas

1. La plupart du temps : une année.

2. Il en existe plusieurs, voir § 1.1

d'un équipement ou aux méthodes de calcul. Evaluer cette courbe amène à estimer de faibles (et fortes) probabilités à ses extrémités, ce qui réclame des échantillons de grandes tailles. Le travail peut s'effectuer par simulations numériques de Monte-Carlo, en propageant les incertitudes grâce à un modèle numérique. La quantité de données nécessaires peut se réduire avec des modèles paramétriques du type log-normal, mais la question de leur représentativité reste ouverte.

Outre les temps de calculs liés à des modèles numériques parfois sophistiqués, la principale difficulté liée à cette problématique provient du manque de données. Si ce problème est inhérent aux méthodes probabilistes en général, il s'accroît ici du fait de la relative rareté des phénomènes sismiques. Les bases de données de signaux sismiques réels sont en effet insuffisamment riches vis-à-vis du nombre de calculs à effectuer. Une étape indispensable consiste donc à enrichir ces bases au moyen de signaux artificiels, générés à partir de modèles supposés représentatifs des scénarios sismiques pour lesquels les structures d'intérêt sont dimensionnées.

Le manque de signaux sismiques réels et la diversité des modèles proposés conduisent, à l'heure actuelle, à l'existence de différentes méthodes d'élaboration des courbes de fragilité. La plus répandue repose sur le *scaling* des accélérogrammes. Elle consiste à générer une famille de signaux artificiels et à la pondérer par différents coefficients croissants, ceci afin d'estimer les probabilités de défaillance correspondantes en calculant les réponses de la structure d'intérêt à l'ensemble de ces signaux. Cette méthode, quoique très répandue, n'est pas cohérente d'un point de vue sismologique : certaines caractéristiques des signaux, comme leur contenu fréquentiel, peuvent en effet évoluer avec leur intensité.

Ce travail de thèse entend développer une méthodologie d'évaluation des courbes de fragilité exploitable en contexte industriel. Elle doit donc impliquer un nombre réduit de simulations numériques de Monte-Carlo, tout en restant au plus proche de la réalité et des données. Nous souhaitons en particulier nous affranchir du *scaling* et de la modélisation log-normale.

Le chapitre 1 de ce mémoire recense les méthodes usuelles de génération de signaux et d'évaluation des courbes de fragilité. Le procédé de simulation retenu dans ce travail fut récemment proposé par Zentner et Poirion [89] [52] [53]. Fondé sur la représentation de Karhunen-Loève des processus aléatoires, il consiste à identifier une telle représentation à l'aide d'une famille de signaux réels de référence, supposés constituer des réalisations d'un même processus³. Ceci permet d'en générer de nouvelles réalisations, aux caractéristiques proches de celles des signaux de référence⁴.

La figure 1 résume notre méthodologie, détaillée au chapitre 2. Elle consiste en premier lieu à identifier une représentation de Karhunen-Loève de l'excitation pour enrichir la base des signaux de référence. Nous proposons ensuite de réaliser un « partage en classes », c'est-à-dire de regrouper les signaux synthétiques par classes croissantes d'indicateur de nocivité sismique. Nous utilisons pour cela un algorithme de partitionnement de données, la méthode des « *k*-moyennes ». Ce type de partage en classes nous permet d'éviter

3. Celui-ci peut éventuellement être non stationnaire et non gaussien.

4. Ceci sera vérifié *a posteriori*.

le *scaling* des accélérogrammes, et donc de préserver les spécificités des signaux dues à leurs intensités. Il présente également l'avantage d'être utilisable quel que soit l'indicateur de nocivité sismique choisi. Mais notre méthode de simulation est également utilisée pour générer, à moindre coût numérique, des réalisations de la réponse de la structure d'intérêt soumise aux signaux sismiques : il suffit pour cela d'identifier une représentation de Karhunen-Loève du processus de réponse. La mise en œuvre industrielle de cette méthode exige par ailleurs de n'utiliser que quelques dizaines de signaux de réponse par classes. La fin du chapitre 2 propose donc une méthodologie pour obtenir, numériquement, un intervalle encerclant la courbe de fragilité, et prenant en compte l'erreur statistique commise sur l'estimation de la fonction de covariance.

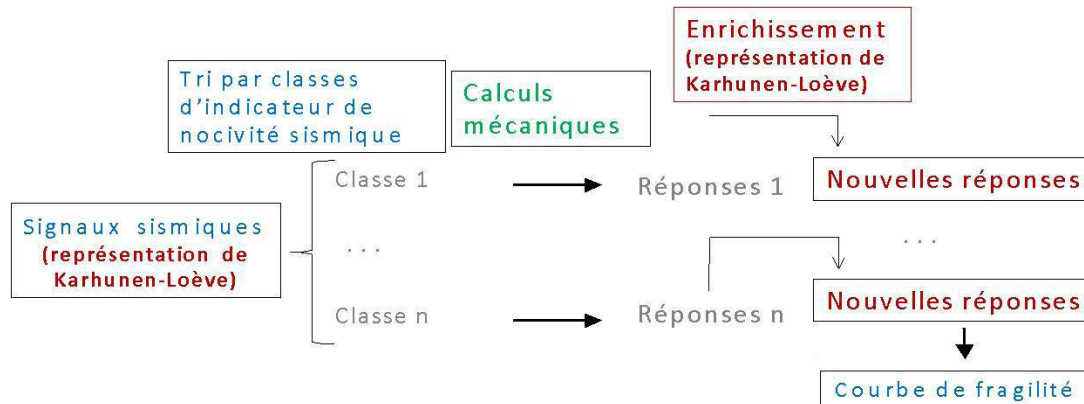


FIGURE 1 – Méthodologie d'obtention des courbes de fragilité fondée sur la représentation de Karhunen-Loève des processus aléatoires.

La validation de l'ensemble de cette méthodologie est alors présentée au chapitre 3, en considérant des structures élémentaires de types oscillateurs linéaire et non-linéaire à un degré de liberté. La mise en œuvre sur un cas « industriel » intervient au chapitre 4, avec un tronçon de la ligne de tuyauterie ASG⁵, dont le modèle éléments-finis fut validé à partir d'essais sur table vibrante au CEA-Saclay.

5. Alimentation de Secours en eau du GV d'un REP 1300.

Chapitre 1

Bibliographie

Le travail de recherche préparatoire a porté sur trois axes principaux :

- la caractérisation du mouvement du sol, avec les notions de nocivité sismique et de spectre de réponse ;
- les différents procédés de simulation de signaux sismiques, jusqu'à celui proposé par Zentner et Poirion [89] [52] [53] ;
- les méthodes habituelles de construction de courbes de fragilité sismique.

1.1 Caractérisation du mouvement du sol

Dans l'ingénierie sismique, le mouvement du sol (le cas échéant dans chaque direction) est généralement caractérisé par son accélération. Celle-ci peut se modéliser par un processus stochastique à valeurs réelles $\{\Gamma_t\}_{t \in [0, T]} = \{\Gamma(t)\}_{t \in [0, T]}$, indexé sur un intervalle de temps $[0, T]$ et défini sur un espace probabilisé (Ω, \mathcal{T}, P) . Il est supposé **centré, du second ordre et continu en moyenne d'ordre deux** (en m.o.d.)¹. Ses réalisations sont appelées **accélérogrammes**, ou **signaux d'entrée**, et sont notés en minuscules : $\{\gamma(t)\}_{t \in [0, T]}$

La **nocivité** d'un accélérogramme $\{\gamma(t)\}_{t \in [0, T]}$, aussi appelée « intensité sismique », se caractérise grâce à divers indicateurs :

- le **PGA** (*Peak Ground Acceleration*), le plus utilisé, défini par :

$$\text{PGA} = \max_{t \in [0, T]} |\gamma(t)|$$

- le **PGV** (*Peak Ground Velocity*) et le **PGD** (*Peak Ground Displacement*) , respectivement vitesse maximale et déplacement maximal du sol, en valeur absolue, sur $[0, T]$;

- le **CAV** (*Cumulative Absolute Velocity*) défini par
$$\text{CAV} = \int_0^T |\gamma(t)| dt$$

1. Des rappels de probabilités réelles sont disponibles dans l'annexe A

— l'**Intensité d'Arias** : $I_{Arias} = \frac{1}{g} \int_0^T \gamma(t)^2 dt$ avec g l'accélération de la pesanteur ;

— l'**Energie** : $E_\gamma(t) = \int_0^t \gamma(\tau)^2 d\tau$, pour $t \in [0, T]$;

— la **durée de phase forte** ou T_{sm} , durée pendant laquelle l'énergie évolue entre 5% et 95% de sa valeur totale. $T_{sm} = t_{95} - t_5$ avec :

- t_5 tel que $E_\gamma(t_5) = 0.05 E_\gamma(T)$
- t_{95} tel que $E_\gamma(t_{95}) = 0.95 E_\gamma(T)$

La **phase forte** est l'intervalle de temps $[t_5, t_{95}]$.

Pour « intégrer » l'effet de la structure, on utilise également le **SRO** (Spectre de Réponse d'Oscillateur) d'un accélérogramme, qui caractérise sa réponse spectrale.

Soit un filtre linéaire à un degré de liberté, de pulsation propre ω_0 et de taux d'amortissement ξ_0 , soumis à l'accélérogramme $\{\gamma(t)\}_{t \in [0, T]}$. Soit $\{y(t)\}_{t \in [0, T]}$ sa réponse en déplacement, qui vérifie donc l'équation différentielle suivante :

$$\ddot{y}(t) + 2\omega_0\xi_0\dot{y}(t) + \omega_0^2y(t) = -\gamma(t)$$

Pour ξ_0 fixé, on nomme **Pseudo-accélération (PSA, Pseudo Spectral Acceleration)** la quantité :

$$PSA(\omega_0) = \omega_0^2 \max_{t \in [0, T]} |y(t)| \stackrel{\text{noté}}{=} \omega_0^2 y_m$$

Alors $\lim_{\omega_0 \rightarrow +\infty} PSA(\omega_0)$ est appelée **ZPA (Zero Period Acceleration)**

$$N.B. : ZPA = \max_{t \in [0, T]} |\gamma(t)| = PGA.$$

Le **SRO de l'accélérogramme** $\{\gamma(t)\}_{t \in [0, T]}$ à **taux d'amortissement** ξ_0 **fixé** représente alors la PSA en fonction de la fréquence $f_0 = \frac{\omega_0}{2\pi}$, comme sur l'exemple de la figure 1.1.

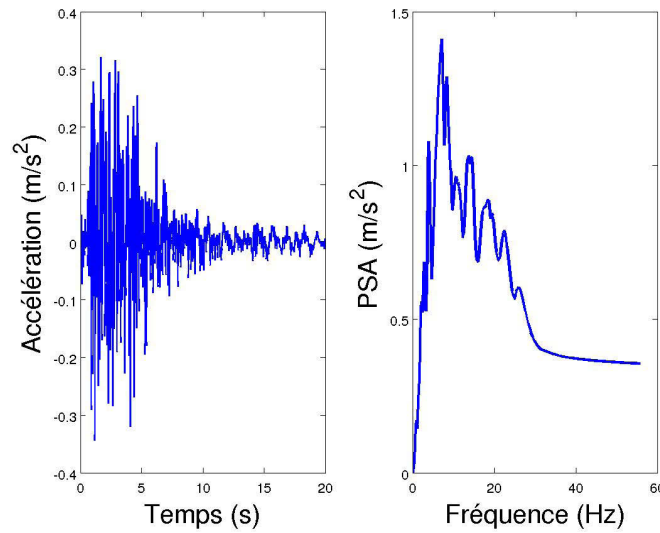


FIGURE 1.1 – Exemple d'un signal réel (à gauche) et de son SRO pour $\xi_0 = 5\%$ (à droite).

Notons que « le PSA d'un accélérogramme » ne signifie rien. Il convient de préciser, soit « le PSA pour un **amortissement** ξ_0 à une **fréquence** f_0 », soit « le Spectre de Réponse pour un **amortissement** ξ_0 ».

1.2 Modèles stochastiques de mouvement sismique

Les courbes de fragilité peuvent s'évaluer par simulation, et la mise en œuvre des calculs de Monte-Carlo réclame de très nombreux accélérogrammes, qui peuvent être soit réels, soit artificiels. Les signaux utilisés sont idéalement réels, proviennent d'un même site sismique et constituent des réalisations du processus stochastique modélisant l'accélération du sol. Comme il est impossible de réunir de tels signaux en très grand nombre, il convient alors de se rabattre sur des accélérogrammes d'origines certes diverses, mais **susceptibles** de constituer des réalisations du processus : ils doivent répondre au même « scénario sismique ». Pour caractériser un scénario sismique, on utilise principalement les deux grandeurs suivantes :

- sa distance épacentrale Δ , représentant la distance de la structure à l'épicentre (ou source) du séisme ;
- sa magnitude M , qui peut revêtir plusieurs significations. La plus simple est la magnitude locale $M_l = \log(A) - \log(A_0) + c \log(\Delta)$ avec A l'amplitude maximale mesurée sur le sismogramme, A_0 une amplitude de référence correspondant à $M_l = 0$ et $\Delta = 100$ km, et c une constante d'étalonnage. Cette équation empirique est inhomogène, et les constantes d'étalonnage A_0 et c ne sont valables que localement. On peut alors définir la magnitude des ondes de surface ou des ondes de volume, qui présentent chacune des défauts ; c'est finalement la complexe magnitude de moment M_ω , non présentée ici, qui est utilisée la plupart du temps aujourd'hui.

Malheureusement, la quantité de signaux compatibles fournis par les bases de données est insuffisante pour mener des études de fiabilité. Les courbes de fragilité s'évaluent donc à l'aide d'accélérogrammes

artificiels, obtenus selon une méthode de génération apte à « représenter » un scénario sismique donné.

Ces différentes méthodes reposent sur deux types de modèles :

- des modèles paramétriques (principalement des bruits blancs filtrés), dont les paramètres sont identifiés à l'aide d'une famille de SRO ou de signaux de référence ;
- des modèles empiriques.

1.2.1 Modèles classiques paramétriques

Rappelons d'abord deux définitions importantes². Un processus $\{X(t)\}_{t \in \mathbb{R}}$ est dit :

- **strictement stationnaire** si $\forall u \in \mathbb{R}$, $\{X(t+u)\}_{t \in \mathbb{R}}$ a même loi que $\{X(t)\}_{t \in \mathbb{R}}$. Sa fonction d'autocorrélation s'écrit dans ce cas $R_X(u) = E[X(t)X(t+u)]$, $\forall t \in \mathbb{R}$;
- **gaussien** si ses lois finies-dimensionnelles sont toutes gaussiennes, *i.e.* si toute combinaison linéaire finie du type $a_1X(t_1) + \dots + a_nX(t_n)$ constitue une v.a. de distribution gaussienne.

Ne perdons pas de vue qu'un processus stationnaire est défini sur \mathbb{R} **tout entier**. Les procédés qui vont suivre s'organisent en deux temps : après avoir modélisé un processus sur \mathbb{R} , il est observé sur un intervalle de temps fini, qui est soit la durée de la phase forte, soit la durée entière du phénomène sismique. La plupart des modèles proposés par les auteurs s'insèrent dans une des catégories suivantes :

1. seule la phase forte du séisme est considérée. Le mouvement du sol est modélisé par un processus stochastique stationnaire gaussien, observé sur l'intervalle de la phase forte ;
2. un processus stationnaire gaussien est observé sur la totalité de la durée du signal, puis « modulé » en temps par une « fonction enveloppe » adéquate ;
3. un processus non stationnaire est observé sur la totalité de la durée du signal, puis modulé ou pas selon les cas.

D'une façon plus précise, la durée de phase forte peut se voir comme une variable aléatoire, dont la distribution est estimée à l'aide des signaux répondant au scénario considéré. Pour générer un nouveau signal, la durée de sa phase forte est simulée avant d'utiliser une des méthodes qui vont suivre.

Modèles stationnaires à DSP

La simulation d'un processus stationnaire gaussien peut s'effectuer grâce à la notion essentielle de **Densité Spectrale de Puissance (DSP)**³. Sous certaines conditions de régularité (précisées par exemple dans [73] et que nous supposons satisfaites ici), un processus stochastique stationnaire $\{X(t)\}_{t \in \mathbb{R}}$ possède une DSP notée S_X , que le théorème de Wiener-Khinchine assure être la transformée de Fourier de la fonction d'autocorrélation R_X du processus. A savoir, $\forall \omega \in \mathbb{R}$:

$$S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_X(\tau) e^{-i\omega\tau} d\tau$$

2. Voir les rappels de probabilités à l'annexe A.

3. Définition précise dans les rappels de probabilités de l'annexe A, p. 91

Une fois connue sa DSP, un processus gaussien stationnaire est aisément simulable grâce au théorème de représentation intégrale (voir par exemple [44], p.10).

Dans beaucoup de situations, le mouvement du sol est fourni à l'ingénieur sous la forme d'un ensemble de SRO réglementaires. Il s'agit donc de simuler à partir de ces données des accélérogrammes artificiels « spectro-compatibles ». Les travaux de Vanmarcke [79] [81] fournissent alors une formule d'estimation de la DSP en fonction des valeurs des SRO. Cette formule a été utilisée sous des formes diversement améliorées par de nombreux auteurs, comme par exemple [55], [9] et [87].

Mais il est possible d'imposer la forme de la DSP, et d'ajuster ses paramètres *a posteriori* par des méthodes statistiques s'appuyant sur la base de signaux réels.

Dans une approche destinée à caractériser analytiquement des ordres de grandeur pour des structures simples, la DSP peut être considérée comme constante, ce qui revient à choisir un bruit blanc⁴ gaussien comme processus d'excitation sismique. Mais la plus célèbre DSP est sans conteste celle proposée par **Kanai-Tajimi** (1960, voir [36] et [78]), qui modélise le mouvement du sol comme un bruit blanc gaussien filtré par un oscillateur linéaire à un degré de liberté. La densité obtenue dépend de trois paramètres : l'intensité du bruit blanc S_0 , la fréquence fondamentale ω_g et l'amortissement ξ_g . Elle s'écrit :

$$S_{KT}(\omega) = S_0 \frac{\omega_g^4 + 4\xi_g^2 \omega_g^2 \omega^2}{(\omega_g^2 - \omega^2)^2 + 4\xi_g^2 \omega_g^2 \omega^2}$$

Le modèle de **Clough et Penzien** (1977, voir [11] ou [46]) n'est qu'un modèle de Kanai-Tajimi filtré. Le bruit blanc d'intensité S_0 traverse deux filtres successifs de paramètres respectifs ω_f , ξ_f , ω_p , ξ_p , ce qui conduit à :

$$S_{CP}(\omega) = H(\omega)S_0H(\omega)^*$$

avec

$$H(\omega) = \frac{-1}{\omega_f^2 - \omega^2 + 2j\xi_f\omega_f} \left(\omega_f^2 + 2j\xi_f\omega_f\omega - (\omega_p^2 + 2j\xi_p\omega_p\omega) \frac{\omega_f^2 + 2j\xi_f\omega_f\omega}{\omega_p^2 - \omega^2 + 2j\xi_p\omega_p} \right)$$

Les paramètres qui régissent ces différents modèles peuvent bien sûr se modéliser par des variables aléatoires. Leurs distributions s'estiment à nouveau à partir de signaux relevant du scénario sismique recherché, à moins qu'elles ne figurent déjà elles-aussi dans la littérature [62]. La simulation d'un signal impose donc au préalable de tirer ces différents paramètres, par exemple par la méthode de l'hypercube latin.

Modulation en temps

Les simulations obtenues avec un des modèles précédents sont ensuite simplement multipliées par une « fonction de modulation » ou « fonction enveloppe » déterministe [57] [2] [64], que nous noterons $\{q(t)\}_{t \in [0, T]}$

4. Voir les rappels de probabilités à l'annexe A.

de façon générique. Il existe plusieurs possibilités pour cette fonction enveloppe. Par exemple, si B et $\lambda > 0$:

$$q(t) = Bt^2 e^{-\lambda t}$$

La fonction de Jennings-Housner [34], très populaire [55], propose un plateau pendant la phase forte. Elle dépend de quatre paramètres t_1 , t_2 , θ et α et s'écrit :

$$q(t) = \left(\frac{t}{t_1}\right)^2 \mathbb{I}_{[0,t_1]} + \mathbb{I}_{[t_1,t_2]} + e^{-\theta(t-t_2)^\alpha} \mathbb{I}_{[t_2,T]}$$

Comme précédemment, les paramètres s'ajustent à l'aide de méthodes statistiques s'appuyant sur la base de signaux réels. Dans le cas où l'on dispose d'une famille de SRO réglementaires, les paramètres peuvent s'ajuster sur les spectres de références, mais une méthode plus complexe consiste à directement ajuster les signaux simulés pour les rendre « spectro-compatibles » [76].

Modèles non-stationnaires

La non-stationnarité des signaux sismiques doit être prise en compte pour plus de réalisme. En tout état de cause, les différents paramètres des modèles doivent là-aussi s'estimer à l'aide d'une famille de signaux.

Une méthode consiste à faire varier la DSP en fonction du temps : c'est la notion de **DSP évolutive** proposée par Priestley [58]. Elle est utilisée par exemple par les articles [1], [43] et [42] qui l'appliquent au modèle de Kanai-Tajimi exposé précédemment. Rezaeian et Der Kiureghian [60], pour leur part, examinent ce modèle de « Kanai-Tajimi évolutif » exclusivement dans le domaine temporel : ils généralisent un modèle plus ancien [49] en considérant sur une durée T :

- un bruit blanc gaussien Ω d'intensité S_0 ;
- la réponse impulsionnelle $h(t; \tau)$ d'un filtre linéaire à 1 degré de liberté (ddl) de paramètres propres évolutifs $\omega_0 - (\omega_0 - \omega_N) \frac{\tau}{T}$ et ξ_f ;
- son écart-type $\sigma_h(t) = \sqrt{\int_{-\infty}^t h(t - \tau; \tau)^2 d\tau}$;
- une fonction de modulation q de paramètres $\hat{\alpha}$;

le processus se construisant alors par convolution :

$$\Gamma(t) = q(t, \hat{\alpha}) \left(\frac{1}{\sigma_h(t)} \int_{-\infty}^t h(t - \tau, \tau) \Omega(\tau) d\tau \right)$$

D'autres méthodes restent également attachées au domaine temporel, comme celle de Sabetta et Pugliese [65] [66]. Certaines, enfin, proposent de filtrer linéairement un train de pulses poissonniens [43], ou encore d'appliquer une transformation non-linéaire à un bruit blanc filtré, construit comme pour un modèle à DSP [83].

1.2.2 Modèle empirique reposant sur la représentation de Karhunen-Loève

La « méthode K-L » est une nouvelle manière de générer des signaux développée par Poirion et Zentner dans leurs articles [89], [52] et [53]. Elle repose sur le théorème de Karhunen-Loève⁵, ainsi nommé en l'honneur du finlandais Kari Karhunen (1915-1992) et du français Michel Loève (1907-1979). Sous certaines conditions de régularité, ce résultat permet de représenter, dans $L^2_{\mathbb{R}}(\Omega, \mathcal{T}, P)$, un processus stochastique centré $\{X(t)\}_{t \in [0, T]}$ par, $\forall t \in [0, T]$:

$$X(t) = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_{\alpha}} \xi_{\alpha} \phi_{\alpha}(t)$$

à l'aide de fonctions $\{\phi_{\alpha}\}_{\alpha \geq 1}$, de coefficients réels $\{\lambda_{\alpha}\}_{\alpha \geq 1}$ et de variables aléatoires réelles $\{\xi_{\alpha}\}_{\alpha \geq 1}$, tous caractérisés par la fonction d'autocorrélation du processus $R_X(t, t') = E[X(t)X(t')]$. Ce résultat, très utile en théorie du signal, a déjà prouvé ses multiples applications dans de nombreuses problématiques industrielles⁶.

Cette méthode sera expliquée en détail au chapitre 2. Elle utilise ce résultat pour, elle aussi, « enrichir » une base d'accélérogrammes réels, supposés, sinon constituer des réalisations d'un même processus stochastique, du moins représenter un même scénario sismique. Après identification des différents paramètres, ces signaux sont « reproduits » et les nouveaux signaux artificiels, censés relever eux-aussi du même scénario, permettent d'engager des calculs de Monte-Carlo.

1.3 Les courbes de fragilité

1.3.1 Définition et généralités

La probabilité de défaillance de la structure, étant donnée l'intensité sismique α à laquelle elle est soumise, est notée $P_f(\alpha)$. La courbe de fragilité consiste donc en le tracé de $P_f(\alpha)$ en fonction de α .

Dans les faits, la défaillance vis-à-vis d'un critère donné est établie lorsqu'une certaine variable d'intérêt relative à l'étanchéité, la stabilité... dépasse un seuil prédéfini, nommé « critère de défaillance ». Dans le cas qui nous occupera, cette variable d'intérêt sera le **maximum du déplacement de la structure**. Ainsi, si la réponse de la structure est représentée par le processus stochastique $\{Y(t)\}_{t \in [0, T]}$, et si le critère de défaillance est noté b , cette probabilité est donnée par :

$$P_f(\alpha) = P\left(\max_{t \in [0, T]} |Y(t)| > b \mid \text{intensité } \alpha\right)$$

Les courbes de fragilité se construisent facilement pour des modèles de structures très simples et de processus d'entrée classiques : bruit blanc gaussien, processus gaussiens, etc ; autant de situations dont le document [20] propose une revue. Il est même parfois possible d'exhiber une forme analytique approchée, grâce aux développements présentés dans des références comme [80] et [56]. Ces cas simples, représentatifs de structures répondant majoritairement sur leur premier mode de vibration sous séisme, ne sont pas que

5. Voir § 2.1.1

6. A titre d'exemple, certains auteurs l'ont utilisé pour modéliser la perméabilité [23] ou la porosité [30].

de simples cas d'école : les solutions de référence qu'ils fournissent sont utiles pour obtenir des ordres de grandeur, indispensables à l'appréhension des cas de structures industrielles complexes.

En dehors de ces cas élémentaires, la construction de la courbe de fragilité peut s'effectuer de différentes façons. Certaines méthodes font l'hypothèse que la courbe suit une loi classique, dont les paramètres doivent être estimés. D'autres ne font aucune hypothèse de forme. Mais dans les deux cas, elles sont menées en contexte Monte-Carlo grâce à une famille d'accélérogrammes obtenus par une méthode de simulation.

1.3.2 Le modèle log-normal

Un des modèles les plus utilisés consiste à supposer que la fonction P_f peut s'assimiler à la fonction de répartition d'une certaine variable aléatoire A , nommée « capacité » [85] :

$$P_f(\alpha) = P(A \leq \alpha)$$

Dans la majorité des cas, cette variable est supposée suivre une distribution log-normale $\mathcal{LN}(A_m, \beta)$ de médiane A_m et d'écart-type logarithmique β :

$$P_f(\alpha) = \int_0^\alpha \frac{1}{t\beta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t/A_m)}{\beta}\right)^2} dt$$

Pour évaluer A_m et β , le secteur nucléaire recourt à un procédé qui fut développé par l'*Electric Power Research Institute* (EPRI) [39]. Cette approche exige une hypothèse simplificatrice supplémentaire. Elle exprime la variable A comme produit d'une capacité de référence A_{SSE} (*Safe Shutdown Earthquake*) et de plusieurs « facteurs de marges » F_i , liées aux différentes sources d'incertitudes du problème traité, comme la résistance mécanique ou l'incohérence du signal sismique :

$$A = A_{SSE} \prod_i F_i$$

Les facteurs de marges F_i , qui sont de nouvelles variables aléatoires, sont supposées être indépendantes et suivre elles-mêmes des distributions log-normales $\mathcal{LN}(F_{i,m}, \beta_{F_i})$, ce qui conserve le caractère log-normal de la variable A . En particulier :

$$A_m = A_{SSE} \prod_i F_{i,m}$$

ce qui signifie finalement que la médiane A_m constitue la marge *best estimate*. Les paramètres de ces facteurs de marges sont tabulés dans les rapports EPRI [18], fruits du retour d'expérience des nombreuses EPS sismiques réalisées aux Etats-Unis depuis 1979.

Une autre possibilité consiste à évaluer directement la médiane A_m et l'écart-type logarithmique β de la variable de capacité en utilisant des méthodes statistiques. La méthode la plus répandue consiste à construire

un modèle de régression linéaire (voire bilinéaire), comme détaillé dans les articles [28], [16], voire [17] pour un raffinement. Les exemples d’application de cette méthode sont légion⁷, le tout pour un coût numérique relativement faible, à savoir moins de cent calculs dans la plupart des cas. Toutefois, la donnée d’un modèle de régression présente l’inconvénient d’introduire une simplification supplémentaire. Il est également possible de recourir à la méthode du maximum de vraisemblance, comme dans [70], [71] et [86]. Son coût numérique est réduit et, contrairement à la méthode précédente, autorise la construction d’intervalles de confiance sans calculs supplémentaires. Enfin, comme les calculs numériques sont très coûteux, l’inférence bayésienne permet de mêler les deux approches précédentes comme illustré dans [77] et [40].

1.3.3 Le *scaling*

Si la majorité des études s’appuient sur le modèle log-normal, il est possible d’estimer la courbe de fragilité d’une structure sans aucune hypothèse préalable sur sa forme, en s’appuyant exclusivement sur une approche Monte-Carlo et sur une (large) famille d’accélérogrammes. La méthode du *scaling* [35] consiste, pour chaque niveau d’intensité sismique requis, à appliquer un coefficient multiplicateur à l’ensemble des accélérogrammes disponibles afin que la moyenne de leurs intensités corresponde désormais au niveau choisi. Ainsi, chaque niveau sismique est représenté par une famille de signaux d’entrée d’intensités adéquates : la proportion de ces signaux conduisant la structure à la défaillance fournit une estimation de la probabilité recherchée.

Cette méthode possède pourtant deux grands points faibles. Son temps de calcul est très élevé, voire prohibitif pour les structures industrielles ; de plus, lorsque le niveau d’intensité sismique augmente, elle suppose que le contenu fréquentiel des accélérogrammes augmente homothétiquement. Cette hypothèse est très simplificatrice comme le montrera le § 3.3. Et, comme pour la méthode log-normale, ce procédé ne peut aboutir, pour des structures linéaires, qu’à des courbes de fragilité strictement croissantes, propriété que rien ne justifie *a priori*.

1.4 Conclusion

Au regard de cette bibliographie, la méthode de génération de signaux proposée par Zentner et Poirion semble la plus appropriée à notre étude. Elle permet en premier lieu de s’affranchir des différentes hypothèses simplificatrices utilisées par les autres méthodes, toutes plus ou moins fondées sur un bruit blanc filtré. Le processus simulé ici conserve toute sa généralité, il n’est ni gaussien ni stationnaire, et n’est construit qu’à partir de la seule base de signaux réels. Elle autorise par ailleurs une nouvelle méthodologie⁸, qui nous est propre, pour construire les courbes de fragilité : sans hypothèse log-normale, et sans avoir recours au principe du *scaling*. La représentativité, la robustesse et l’efficacité de cette méthode de simulation seront validées sur des modèles simples de structures au chapitre 3, avant d’être testée sur un exemple concret de structure industrielle au chapitre 4.

7. Vieux immeubles du pourtour de la Méditerranée [27], ponts routiers [38], etc.

8. Fondée sur un partitionnement de données, elle est détaillée au § 2.4.

Chapitre 2

Méthodologie

L'évaluation d'une courbe de fragilité avec la méthode K-L se déroule en deux temps. Une famille de signaux d'entrée réels, choisis pour leur adéquation à un scénario sismique donné, est « enrichie ». Ceci nécessite de simuler des variables aléatoires **dépendantes** à l'aide d'un certain nombre de leurs réalisations. Ces signaux d'entrée synthétiques sont ensuite triés par classe d'indicateur de nocivité sismique, puis confrontés au modèle numérique de la structure étudiée. Les signaux de réponse qui en découlent peuvent soit être utilisés comme tels, soit être eux-mêmes enrichis.

Cette méthodologie ne dispense pas de prendre en compte les incertitudes, qui sont nombreuses. Le paragraphe 2.5 recense les incertitudes épistémique et statistique, ainsi que celles pesant sur le modèle K-L. L'incertitude liée au modèle numérique sera quant à elle négligée.

2.1 Enrichissement d'une famille de signaux : « méthode K- L »

L'article [53] de Zentner et Poirion propose une nouvelle méthode, fondée sur le théorème de Karhunen-Loève [37] [45], pour simuler un processus **non gaussien et non stationnaire**. Sur l'hypothèse que l'accélération du sol peut se modéliser par un processus stochastique $\{\Gamma(t)\}_{t \in [0, T]}$, elle utilise une base d'accélérogrammes réels, **sensés constituer des réalisations du processus Γ** , pour en simuler de nouvelles.

Dans tout ce qui suit, H désigne l'espace de Hilbert $L^2_{\mathbb{R}}([0, T], dt)$, muni du produit scalaire usuel :

$$(f|g) = \int_0^T f(t)g(t) dt \quad (2.1)$$

La norme préhilbertienne associée est alors :

$$\|f\| = \sqrt{(f|f)} = \sqrt{\int_0^T f(t)^2 dt} \quad (2.2)$$

2.1.1 Une forme du théorème de Karhunen-Loève

Une démonstration détaillée de ce résultat est proposée en annexe C.

Soit $X = \{X(t)\}_{t \in [0, T]}$ un processus stochastique réel défini sur un espace probabilisé (Ω, \mathcal{T}, P) et indexé sur $[0, T]$, du second ordre, centré, continu en m.o.d., à trajectoires continues¹.

Soit $R_X : (t, t') \rightarrow E[X(t)X(t')]$ la fonction d'autocorrélation du processus X .

L'opérateur Q_X est défini comme :

$$Q_X : \begin{cases} H & \rightarrow H \\ f & \rightarrow \begin{cases} [0, T] & \rightarrow \mathbb{R} \\ t & \rightarrow \int_0^T R_X(t, t') f(t') dt' \end{cases} \end{cases}$$

La résolution du problème aux valeurs propres $Q_X(\phi) = \lambda\phi$ conduit à l'existence de :

- $\{\lambda_\alpha\}_{\alpha \geq 1}$ les valeurs propres de Q_X , une famille de réels positifs, ordonnés de façon décroissante ;
- $\{\phi_\alpha\}_{\alpha \geq 1}$ des fonctions propres associées aux valeurs propres $\{\lambda_\alpha\}_{\alpha \geq 1}$, formant une base hilbertienne de $(H, (\cdot, \cdot))$.

De plus, il existe $\{\xi_\alpha\}_{\alpha \geq 1}$ une famille dénombrable de v.a. réelles définies sur (Ω, \mathcal{T}, P) :

- **centrées et de variance 1** ;
- **deux à deux orthogonales** (*i.e.* non corrélées) et *a priori* dépendantes car non gaussiennes dans le cas général ;

vérifiant, $\forall t \in [0, T]$:

$$X(t) = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t) \quad (2.3)$$

Cette identité est connue sous le nom de **représentation de Karhunen-Loève**. La convergence y est uniforme en moyenne quadratique, et les v.a. $\{\xi_\alpha\}_{\alpha \geq 1}$ s'obtiennent selon :

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \int_0^T X(t) \phi_\alpha(t) dt \quad (2.4)$$

2.1.2 Identification de l'excitation : « modèle K-L »

Fort de ce résultat, nous avons retenu une représentation du même type pour le processus d'excitation sismique $\Gamma = \{\Gamma(t)\}_{t \in [0, T]}$, en tronquant la somme (2.3) à un rang M appelé **rang de troncature**. Avant toute chose il convient de substituer au processus Γ un processus X_Γ satisfaisant aux hypothèses du théorème de Karhunen-Loève. Comme les réalisations de Γ sont disponibles sous forme échantillonnée², il est toujours

1. Voir définitions en annexe A.

2. Voir les précisions sur l'échantillonnage à l'annexe B.

possible de le considérer comme un processus du second ordre, continu en m.o.d. et à trajectoires continues. Mais il n'est pas nécessairement centré, et doit céder la place au nouveau processus $X_\Gamma = \{X_\Gamma(t)\}_{t \in [0, T]}$ vérifiant $\forall t \in [0, T]$:

$$X_\Gamma(t) = \Gamma(t) - E[\Gamma(t)] \quad (2.5)$$

Après le choix d'un rang de troncature M , $X_\Gamma(t)$ est estimé $\forall t \in [0, T]$ par la somme suivante :

$$\bar{X}_\Gamma(t) = \sum_{\alpha=1}^M \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t) \quad (2.6)$$

Le **modèle K-L** consiste alors à estimer Γ par $\bar{\Gamma}$ vérifiant $\forall t \in [0, T]$:

$$\boxed{\bar{\Gamma}(t) = E[\Gamma(t)] + \sum_{\alpha=1}^M \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t)} \quad (2.7)$$

avec :

- $\{\xi_\alpha\}_{1 \leq \alpha \leq M}$ des v.a. réelles du second ordre, **décorrélées, centrées et de variance 1**, données par :

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \int_0^T X_\Gamma(t) \phi_\alpha(t) dt \quad (2.8)$$

- $\{(\lambda_\alpha, \phi_\alpha)\}_{1 \leq \alpha \leq M}$ les M premiers éléments propres (ordonnés par valeurs décroissantes des λ_α) de l'opérateur Q_{X_Γ} .

Le rang de troncature M détermine la proportion d'énergie de X_Γ conservée. En effet, puisque les ξ_α sont centrées et de variance 1 :

$$\|X_\Gamma\|^2 = \left(\sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha \middle| \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha \right) = \sum_{\alpha=1}^{+\infty} \lambda_\alpha \|\phi_\alpha\|^2 \quad (2.9)$$

et :

$$\|\bar{X}_\Gamma\|^2 = \left(\sum_{\alpha=1}^M \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha \middle| \sum_{\alpha=1}^M \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha \right) = \sum_{\alpha=1}^M \lambda_\alpha \|\phi_\alpha\|^2 \quad (2.10)$$

Si l'on souhaite ainsi conserver une proportion p d'énergie, M est choisi comme le plus grand rang pour lequel $\frac{\|\bar{X}_\Gamma\|^2}{\|X_\Gamma\|^2} \leq p$.

Dans l'intégralité de ce mémoire, $p = 99\%$.

Ce modèle est identifié à l'aide des signaux sismiques disponibles en utilisant la procédure suivante.

En pratique³, l'intervalle $[0, T]$ est échantillonné en une $(N-1)$ -partition régulière

$$0 = t_1 < t_2 < \dots < t_N = T \quad \text{de pas} \quad \Delta t = \frac{T}{N-1}$$

La famille de signaux réels dont nous disposons est supposée s'écrire $\{\{\gamma^{(l)}(t)\}_{t \in [0, T]}\}_{1 \leq l \leq L}$. Chaque $\gamma^{(l)}$ est fourni sous la forme d'une famille finie de réels $\{\gamma^{(l)}(t_j)\}_{1 \leq j \leq N}$ correspondant à ses valeurs enregistrées aux nœuds $\{t_j\}_{1 \leq j \leq N}$.

On dispose ainsi d'un L -échantillon numérique de Γ , c'est-à-dire d'un L -échantillon numérique $\{\{x_\Gamma^{(l)}(t_j)\}_{1 \leq j \leq N}\}_{1 \leq l \leq L}$ de X_Γ .

A partir de cet échantillon, la fonction d'autocorrélation R_{X_Γ} est estimée par l'estimateur statistique standard $\bar{R}_{X_\Gamma} \in \mathcal{M}_N(\mathbb{R})$ vérifiant, $\forall (i, j) \in \llbracket 1; N \rrbracket^2$:

$$[\bar{R}_{X_\Gamma}]_{ij} = \frac{1}{L-1} \sum_{l=1}^L x_\Gamma^{(l)}(t_i) x_\Gamma^{(l)}(t_j) \quad (2.11)$$

Portant cette estimée dans l'opérateur \mathcal{Q}_{X_Γ} , le problème est ensuite résolu numériquement par une méthode de Galerkin [10]. Soient $\{(\bar{\lambda}_\alpha, \{\bar{\phi}_\alpha(t_j)\}_{1 \leq j \leq N})\}_{1 \leq \alpha \leq M}$ les M couples-solutions obtenus.

En introduisant ces couples ainsi que les réalisations $x_\Gamma^{(l)}$ dans l'intégrale (2.8), que l'on approxime par une méthode de rectangles à gauche sur la partition considérée de $[0, T]$, on obtient pour chaque v.a. ξ_α un L -échantillon de réalisations approchées $\{\tilde{\xi}_\alpha^{(l)}\}_{1 \leq l \leq L}$ vérifiant, $\forall (l, \alpha) \in \llbracket 1; L \rrbracket \times \llbracket 1; M \rrbracket$:

$$\tilde{\xi}_\alpha^{(l)} = \frac{1}{\sqrt{\bar{\lambda}_\alpha}} \sum_{j=1}^N x_\Gamma^{(l)}(t_j) \bar{\phi}_\alpha(t_j) \Delta t \quad (2.12)$$

Dans une première approximation, nous pourrions supposer que les v.a. $\{\xi_\alpha\}_{1 \leq \alpha \leq M}$ sont indépendantes. Il suffirait alors d'estimer leurs lois marginales pour simuler chacune de ces variables, et par suite le processus Γ lui-même, grâce à (2.7). Mais nous verrons au § 3.4 qu'une telle hypothèse conduit à de moins bons résultats. Leur dépendance doit donc être prise en compte, par le biais de leur densité de probabilité conjointe, comme détaillé ci-après.

2.2 Estimation d'une densité conjointe

La dépendance de variables $\{\xi_\alpha\}_{\alpha \in \llbracket 1; M \rrbracket}$ est prise en compte grâce à la densité de probabilité conjointe⁴ p_ξ du vecteur aléatoire $\xi = (\xi_1, \dots, \xi_M)$ vérifiant, pour tout borélien B de \mathbb{R}^M :

$$P(\xi \in B) = \int_B p_\xi(x) dx \quad (2.13)$$

3. Voir notations de l'annexe B

4. Voir les rappels sur les vecteurs aléatoires p. 88

Rappelons que cette densité conjointe intervient dans l'écriture de la probabilité marginale de chacune des variables, soit $\forall i \in \llbracket 1; M \rrbracket$ et $x \in \mathbb{R}$:

$$p_{\xi_i}(x) = \int_{\mathbb{R}^{M-1}} p_{\xi}(t_1, \dots, t_{i-1}, x, t_{i+1}, \dots, t_M) dt_1 \dots dt_{i-1} dt_{i+1} \dots dt_M \quad (2.14)$$

Lorsque des v.a. $\{\xi_{\alpha}\}_{1 \leq \alpha \leq M}$ sont disponibles sous la forme de L -échantillons de réalisations $\{\tilde{\xi}_{\alpha}^{(l)}\}_{1 \leq l \leq L}$, ce qui revient à disposer de réalisations $\{\tilde{\xi}^{(l)}\}_{1 \leq l \leq L}$ du vecteur ξ , l'estimation de la densité conjointe p_{ξ} est possible mais reste un problème en général délicat. Nous avons exploré deux pistes différentes : une première qui nous est propre, fondée sur un développement en série, et une seconde, classique, utilisant des fonctions noyaux.

2.2.1 Développement sur une base de polynômes orthogonaux

Notations et rappels sur les polynômes de Hermite

Dans le cadre de la notation multi-indice, si $\alpha = (\alpha_1, \dots, \alpha_i, \dots) \in \mathbb{N}^{\mathbb{N}}$, on note $|\alpha| = \sum_{k=1}^{+\infty} \alpha_k$.

Si α et $\beta \in \mathcal{J} = \{\alpha \in \mathbb{N}^{\mathbb{N}}; |\alpha| < +\infty\}$, on note $\alpha! = \prod_{k=1}^{+\infty} \alpha_k!$ et $\delta_{\alpha\beta} = \prod_{k=1}^{+\infty} \delta_{\alpha_k\beta_k}$

Si les valeurs des composantes de α l'autorisent, on peut définir $\alpha - 1 = (\alpha_1 - 1, \dots, \alpha_i - 1, \dots)$.

Soit $M \geq 2$.

\mathbb{R}^M est muni de sa tribu de Lebesgue $B_0(\mathbb{R}^M)$, la complétée de sa tribu borélienne B_M pour la mesure de Lebesgue $d\mathbf{x}$ (restreinte aux boréliens) [59][63]. L'espace euclidien $(\mathbb{R}^M, \|\cdot\|)$ est alors muni de sa mesure gaussienne canonique définie, $\forall A \in B_0(\mathbb{R}^M)$, par :

$$d\gamma_M(A) = \int_A \frac{1}{(2\pi)^{M/2}} e^{-\frac{\|\mathbf{x}\|_2^2}{2}} d\mathbf{x} \quad (2.15)$$

La densité de la mesure gaussienne canonique sur $(\mathbb{R}^M, B_0(\mathbb{R}^M))$ est donc définie, $\forall \mathbf{x} \in \mathbb{R}^M$, par :

$$\nu_M(\mathbf{x}) = \frac{1}{(2\pi)^{M/2}} e^{-\frac{\|\mathbf{x}\|_2^2}{2}} \quad (2.16)$$

Dans tout ce qui suit, H_{γ_M} désignera l'espace de Hilbert $L_{\mathbb{R}}^2(\mathbb{R}^M, d\gamma_M)$ des (classes modulo les fonctions nulles $d\gamma_M$ -presque partout de) fonctions de \mathbb{R}^M dans \mathbb{R} de carré intégrable par rapport à la mesure gaussienne canonique $d\gamma_M$, cet espace étant muni du nouveau produit scalaire :

$$(f|g)_{H_{\gamma_M}} = \int_{\mathbb{R}^M} f(\mathbf{x})g(\mathbf{x}) d\gamma_M(\mathbf{x}) = \int_{\mathbb{R}^M} f(\mathbf{x})g(\mathbf{x}) \nu_M(\mathbf{x}) d\mathbf{x} \quad (2.17)$$

Soit $g : \mathbb{R} \rightarrow \mathbb{R} : t \rightarrow e^{-t^2/2}$

Dans la pratique, on identifie les vecteurs $(t_1, \dots, t_N, 0, \dots, 0, \dots)$ et (t_1, \dots, t_N) . Rappelons que :

- si $n \in \mathbb{N}$, le n^e polynôme de Hermite sur \mathbb{R} est défini comme $H_n(t) = (-1)^n e^{t^2/2} g^{(n)}(t)$
- si $\alpha \in \mathcal{J}$, on note $H_\alpha = \bigotimes_{k=1}^{+\infty} H_{\alpha_k}$, c'est-à-dire que : $H_{(\alpha_1, \dots, \alpha_i, \dots)}(t_1, \dots, t_i, \dots) = \prod_{k=1}^{+\infty} H_{\alpha_k}(t_k)$
- et les polynômes normalisés de Hermite sont pour leur part définis comme $h_\alpha = \frac{1}{\sqrt{\alpha!}} H_\alpha$.

La famille $(h_\alpha)_{\alpha \in \mathcal{J}}$ constitue une base hilbertienne de $(H_{\gamma_N}, (\cdot | \cdot)_{H_{\gamma_N}})$ [8].

Mise en œuvre

Notre première idée consistait à écrire la densité de probabilité conjointe p_ξ de manière exacte, comme une somme infinie. Pour son estimation, seuls quelques termes de la somme sont alors retenus, et ses coefficients sont évalués grâce aux réalisations $\{\tilde{\xi}^{(l)}\}_{1 \leq l \leq L}$ du vecteur ξ .

Ici, M correspond au nombre de modes retenus dans la représentation de Karhunen-Loève.

Soit \mathcal{G}_M le vecteur multinormale de loi $\mathcal{N}(\mathbf{0}, I_M)$.

Pour la suite, il est indispensable que $\frac{p_\xi}{v_M} \in H_{\gamma_M}$, ce qui est le cas sous la condition suffisante que $\frac{p_\xi}{v_M}$ soit majoré. Mais cette dernière condition peut toujours être considérée comme vérifiée. En effet, comme p_ξ provient d'estimations discrètes et quitte à interpoler, cette fonction peut sans problème être supposée continue et à support compact $K = [-c, c]^M$, ce qui fournit $\forall u \in \mathbb{R}^M$:

$$\begin{aligned} \left| \frac{p_\xi(u)}{v_M(u)} \right| &\leq |p_\xi(u)| (2\pi)^{M/2} e^{-\frac{\|\mathbf{x}\|_2^2}{2}} \\ &\leq \|p_\xi\|_{\infty, K} \times (2\pi)^{\frac{M}{2}} e^{-\frac{Mc^2}{2}} \end{aligned}$$

Comme $(h_\alpha)_{\alpha \in \mathcal{J}}$ est une base hilbertienne de $(H_{\gamma_M}, (\cdot | \cdot)_{H_{\gamma_M}})$, il existe des coefficients $\{q_\alpha\}_{\alpha \in \mathcal{J}}$ tels que :

$$\frac{p_\xi}{v_M} = \sum_{|\alpha|=0}^{+\infty} q_\alpha h_\alpha \quad (2.18)$$

ce qui implique $\forall \alpha \in \mathcal{J}$:

$$\begin{aligned} q_\alpha &= \left(\frac{p_\xi}{v_M} \middle| h_\alpha \right)_{H_{\gamma_M}} \\ &= \int_{\mathbb{R}^M} \frac{p_\xi(u)}{v_M(u)} h_\alpha(u) v_M(u) du \\ &= \int_{\mathbb{R}^M} p_\xi(u) h_\alpha(u) du \\ &= E[h_\alpha(\xi)] \text{ d'après le théorème de transfert} \end{aligned} \quad (2.19)$$

Ainsi :

$$p_{\xi} = \nu_M \sum_{|\alpha|=0}^{+\infty} E[h_{\alpha}(\xi)] h_{\alpha} = \nu_M \sum_{|\alpha|=0}^{+\infty} \frac{1}{\alpha!} E[H_{\alpha}(\xi)] H_{\alpha} \quad (2.20)$$

ou

$$p_{\xi} = \nu_M (S_r + R_r) \quad (2.21)$$

en notant, $\forall r \in \mathbb{N}$:

$$S_r : \mathbb{R}^M \rightarrow \mathbb{R} : u \rightarrow \sum_{|\alpha|=0}^r \frac{1}{\alpha!} E[H_{\alpha}(\xi)] H_{\alpha} \quad (2.22)$$

$$R_r : \mathbb{R}^M \rightarrow \mathbb{R} : u \rightarrow \sum_{|\alpha|=r+1}^{+\infty} \frac{1}{\alpha!} E[H_{\alpha}(\xi)] H_{\alpha} \quad (2.23)$$

En s'appuyant sur les réalisations disponibles $\{\tilde{\xi}^{(l)}\}_{1 \leq l \leq L}$ du vecteur ξ , on peut facilement estimer $E[H_{\alpha}(\xi)]$ par :

$$\overline{H_{\alpha}(\xi)} = \frac{1}{L} \sum_{l=1}^L H_{\alpha}(\tilde{\xi}^{(l)}) \quad (2.24)$$

Ainsi, pour estimer la densité de probabilité conjointe p_{ξ} , nous proposons, pour chaque rang de troncature $r \in \mathbb{N}$, l'estimateur $\overline{p_{\xi}}$ vérifiant :

$$\overline{p_{\xi}} = \nu_M \sum_{|\alpha|=0}^r \frac{1}{\alpha!} \overline{H_{\alpha}(\xi)} H_{\alpha} \quad (2.25)$$

2.2.2 L'estimateur à noyaux

Pour ne pas s'encombrer inutilement de notations supplémentaires, ce second estimateur sera à nouveau noté $\overline{p_{\xi}}$.

Une seconde possibilité consiste à utiliser la méthode KDE (*Kernel Density Estimation* [72], idéale pour les v.a. continues), qui propose d'estimer p_{ξ} , pour tout $\mathbf{x} = (x_1, \dots, x_M) \in \mathbb{R}^M$, par :

$$\overline{p_{\xi}}(\mathbf{x}) = \frac{1}{L h_1 \dots h_M} \sum_{l=1}^L \prod_{\alpha=1}^M K_{\alpha} \left(\frac{x_{\alpha} - \tilde{\xi}_{\alpha}^{(l)}}{h_l} \right) \quad (2.26)$$

Les h_l sont des largeurs à optimiser et les K_{α} des fonctions noyaux. Chaque K_{α} doit être :

- paire ;
- intégrable, vérifiant $\int_{\mathbb{R}} K_{\alpha}(t) dt = 1$ et $\int_{\mathbb{R}} t^2 K_{\alpha}(t) dt \neq 0$;
- de carré intégrable.

La forme de ces noyaux importe moins que les valeurs des paramètres h_l : trop petits, ils rendront le résultat sensible au bruit ; trop grands, certains comportements essentiels, comme la multimodalité, risqueraient de disparaître. Pour simplifier, les coefficients h_l sont choisis tous égaux à une unique valeur h , et les noyaux K_{α} identiques à un noyau K supposé gaussien.

L'estimation \overline{p}_ξ de la densité p_ξ s'écrit ainsi :

$$\overline{p}_\xi(\mathbf{x}) = \frac{1}{Lh^M} \sum_{l=1}^L \prod_{\alpha=1}^M \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{|x_\alpha - \tilde{\xi}_\alpha^{(l)}|}{h} \right)^2} \quad (2.27)$$

soit :

$$\overline{p}_\xi(\mathbf{x}) = \sum_{l=1}^L \frac{1}{L} \frac{e^{-\frac{\|\mathbf{x} - \tilde{\xi}^{(l)}\|_2^2}{2h^2}}}{\sqrt{2\pi}^M h^M} \quad (2.28)$$

La valeur optimale de h est celle qui minimise l'erreur quadratique moyenne. L'approximation de Silverman⁵ [72] propose d'utiliser :

$$h = \left(\frac{4\hat{\sigma}^5}{3L} \right)^{1/5} \simeq 1,06 \hat{\sigma} L^{-1/5} \quad (2.29)$$

avec, dans le cas de ce travail, la variance $\hat{\sigma} = 1$.

Comparaison et conclusion

Par souci de clarté, ces deux méthodes sont testées en dimension un.

Les réalisations $\tilde{\xi}_\alpha^{(l)}$ utilisées proviennent d'une variable particulière⁶, intervenant dans l'étape de validation du chapitre 3. Elle a été choisie pour son fort caractère « pathologique ». La figure 2.1 compare les deux estimations de la densité de probabilité de cette variable :

- avec les polynômes de Hermite (en pointillés bleus), pour lesquels un certain nombre de termes de la somme ont été retenus : 10 à gauche, puis 60 à droite ;
- à l'aide de l'estimateur à noyaux (en traits pleins rouges).

5. ou « Rule of the thumb »...

6. Variable issue du troisième mode de la cinquième classe des 10 000 signaux de réponse présentés au § 3.4.

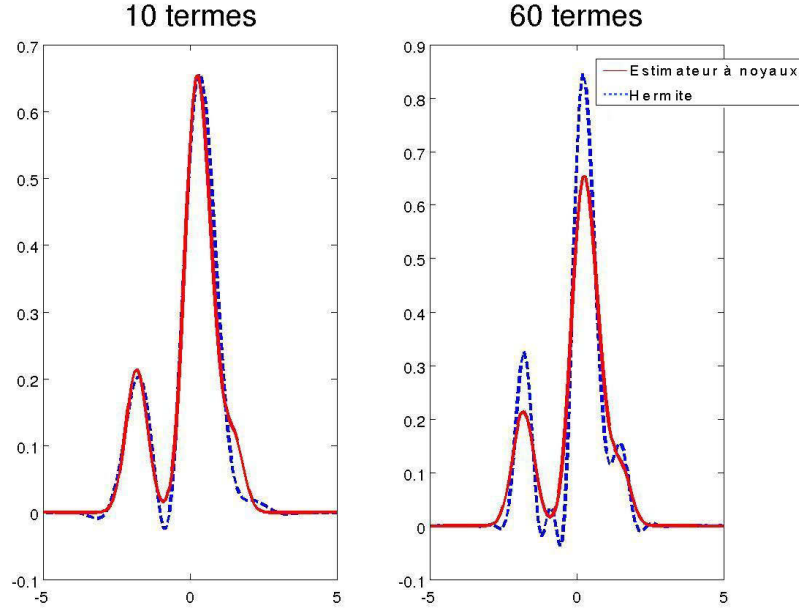


FIGURE 2.1 – Exemples d’estimations, selon les méthodes, de densités de probabilité d’une variable aléatoire.

Dans notre exemple, une adéquation est atteinte au rang 10. Mais la densité estimée avec la méthode hilbertienne prend des valeurs négatives, ce qui n’est pas acceptable, et le phénomène s’amplifie lorsque le rang de troncature augmente. Au vu de l’équation (2.20), il conviendrait peut-être d’effectuer la troncature non pas sur la valeur de α , mais sur le degré des monômes retenus ; calcul qui s’annonce fastidieux. Voilà un point susceptible d’être exploré plus avant après ce travail de thèse.

Nous choisissons d’utiliser l’estimateur à noyaux dans tout ce qui va suivre.

2.3 Simulation de variables dépendantes

Au vu des résultats du paragraphe précédent, l’estimation $\overline{p_{\xi}}$ de la densité p_{ξ} d’un vecteur ξ s’obtient désormais selon la méthode de l’estimateur à noyaux.

Soit $\underline{\xi} = (\xi_1, \dots, \xi_M)$ le vecteur aléatoire de \mathbb{R}^M de densité $p_{\underline{\xi}} = \overline{p_{\underline{\xi}}}$. Nous cherchons à présent à simuler $\underline{\xi}$ comme approximation de ξ . Une première méthode est proposée par l’article [54], une seconde nous est personnelle.

Simulation directe

Pour simuler le vecteur $\underline{\xi}$ de densité conjointe $p_{\underline{\xi}}$, Poirion et Zentner proposent dans l’article [54] de considérer :

- $U \hookrightarrow \mathcal{U}(1, L)$ une variable de loi uniforme discrète : $\forall l \in \llbracket 1; L \rrbracket \quad P(U = l) = \frac{1}{L}$
- $\forall l \in \llbracket 1; L \rrbracket, \mathbf{Y}_l \hookrightarrow \mathcal{N}(\tilde{\xi}^{(l)}, h^2 I_M)$, vecteur aléatoire de loi multivariée d'espérance $\tilde{\xi}^{(l)}$ et de variance-covariance la matrice scalaire $h^2 I_M$. Sa densité de probabilité conjointe est donc $\forall \mathbf{x} \in \mathbb{R}^M$:

$$p_{Y_l}(\mathbf{x}) = \frac{e^{-\frac{\|\mathbf{x} - \tilde{\xi}^{(l)}\|^2}{2h^2}}}{\sqrt{2\pi^M} h^M} \quad (2.30)$$

L'article [54] assure que le vecteur aléatoire \mathbf{Y}_U , indexé par U , et $\underline{\xi}$ sont égaux en loi ; résultat dont voici toutefois une démonstration précise. Pour tout borélien \mathbf{B} de \mathbb{R}^M :

$$(\mathbf{Y}_U \in \mathbf{B}) = \bigcup_{l=1}^L (U = l) \cap (\mathbf{Y}_U \in \mathbf{B}) \text{ réunion d'événements incompatibles.} \quad (2.31)$$

Alors, d'après la σ -additivité des probabilités :

$$P(\mathbf{Y}_U \in \mathbf{B}) = \sum_{l=1}^L P((U = l) \cap (\mathbf{Y}_U \in \mathbf{B})) \quad (2.32)$$

Il vient alors que :

$$\begin{aligned} P(\mathbf{Y}_U \in \mathbf{B}) &= \sum_{l=1}^L P(U = l) P(\mathbf{Y}_U \in \mathbf{B} | U = l) \\ &= \sum_{l=1}^L P(U = l) P(\mathbf{Y}_l \in \mathbf{B}) \\ &= \sum_{l=1}^L P(U = l) \int_{\mathbf{B}} p_{Y_l}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbf{B}} \left(\sum_{l=1}^L P(U = l) p_{Y_l}(\mathbf{x}) \right) d\mathbf{x} \\ &= \int_{\mathbf{B}} \sum_{l=1}^L \frac{1}{L} \frac{e^{-\frac{\|\mathbf{x} - \tilde{\xi}^{(l)}\|_2^2}{2h^2}}}{\sqrt{2\pi^M} h^M} d\mathbf{x} \\ &= \int_{\mathbf{B}} \overline{p_{\xi}}(\mathbf{x}) d\mathbf{x} \\ &= P(\underline{\xi} \in \mathbf{B}) \end{aligned} \quad (2.33)$$

Ainsi $\boxed{\underline{\xi} \stackrel{f}{=} \mathbf{Y}_U}$ (égalité des variables en loi). Comme U et \mathbf{Y}_U sont simples à simuler, il est facile de générer des réalisations $\{\tilde{\xi}^{(k)}\}_{k \in \llbracket 1; n \rrbracket}$ du vecteur $\underline{\xi}$, c'est-à-dire des réalisations **dépendantes** des variables ξ_1, \dots, ξ_M ,

en respectant l'algorithme suivant :

1. la variable discrète uniforme U est simulée ; soit $l \in \llbracket 1; L \rrbracket$ une telle réalisation ;
2. le vecteur gaussien Y_l est simulé. Soit $y_l^{(1)}$ une telle réalisation ;
3. on pose $\underline{\tilde{\xi}}^{(1)} = (\tilde{\xi}_1^{(1)}, \dots, \tilde{\xi}_M^{(1)}) = y_l^{(1)}$, on retourne à l'étape 1 pour simuler un second vecteur, et ainsi de suite.

Variations sur l'échantillonneur de Gibbs

Proposons à présent une seconde méthode, qui est nous est propre, pour obtenir n réalisations $\{\underline{\tilde{\xi}}^{(k)}\}_{k \in \llbracket 1; n \rrbracket}$ de $\underline{\xi}$. Elle consiste à entremêler le procédé précédent avec le traditionnel échantillonneur de Gibbs [22].

Soit $\underline{\tilde{\xi}}^{(0)} \in \mathbb{R}^M$ un vecteur initial donné.

Pour $k \in \llbracket 0; n-1 \rrbracket$, on suppose $\underline{\tilde{\xi}}^{(k)} \in \mathbb{R}^M$ déjà simulé.

Alors $\underline{\tilde{\xi}}^{(k+1)} = (\tilde{\xi}_1^{(k+1)}, \dots, \tilde{\xi}_M^{(k+1)})$ est obtenu en générant **successivement**, $\forall i \in \llbracket 1; M \rrbracket$, chacun des $\tilde{\xi}_i^{(k+1)}$ selon la densité :

$$q_i = p_{\tilde{\xi}_i} \left(\cdot | (\xi_1 = \tilde{\xi}_1^{(k+1)}) \cap \dots \cap (\xi_{i-1} = \tilde{\xi}_{i-1}^{(k+1)}) \cap (\xi_{i+1} = \tilde{\xi}_{i+1}^{(k)}) \cap \dots \cap (\xi_M = \tilde{\xi}_M^{(k)}) \right) \quad (2.34)$$

Soit Z_i la v.a. de \mathbb{R} de densité $p_{Z_i} = q_i$. Alors $\forall x_i \in \mathbb{R}$:

$$\begin{aligned} p_{Z_i}(x_i) &= p_{\tilde{\xi}_i} \left(x_i | (\xi_1 = \tilde{\xi}_1^{(k+1)}) \cap \dots \cap (\xi_{i-1} = \tilde{\xi}_{i-1}^{(k+1)}) \cap (\xi_{i+1} = \tilde{\xi}_{i+1}^{(k)}) \cap \dots \cap (\xi_M = \tilde{\xi}_M^{(k)}) \right) \\ &= \frac{p_{\tilde{\xi}}(\tilde{\xi}_1^{(k+1)}, \dots, \tilde{\xi}_{i-1}^{(k+1)}, x_i, \tilde{\xi}_{i+1}^{(k)}, \dots, \tilde{\xi}_M^{(k)})}{p_{(\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_M)}(\tilde{\xi}_1^{(k+1)}, \dots, \tilde{\xi}_{i-1}^{(k+1)}, \tilde{\xi}_{i+1}^{(k)}, \dots, \tilde{\xi}_M^{(k)})} \\ &\stackrel{\text{noté}}{=} \frac{A}{B} \end{aligned} \quad (2.35)$$

D'une part :

$$\begin{aligned}
A &= \sum_{l=1}^L \frac{1}{L} \left(\prod_{j=1}^{i-1} e^{-\frac{|\tilde{\xi}_j^{(k+1)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right) \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi h}} \left(\prod_{j=i+1}^M e^{-\frac{|\tilde{\xi}_j^{(k)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right) \\
&= \sum_{l=1}^L \frac{1}{L} \left(\prod_{j=1}^{i-1} e^{-\frac{|\tilde{\xi}_j^{(k+1)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \times \prod_{j=i+1}^M e^{-\frac{|\tilde{\xi}_j^{(k)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right) \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi h}} \\
&= \sum_{l=1}^L \frac{1}{L} \beta_l \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi h}}
\end{aligned} \tag{2.36}$$

après avoir posé $\forall l$:

$$\beta_l = \left(\prod_{j=1}^{i-1} e^{-\frac{|\tilde{\xi}_j^{(k+1)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \times \prod_{j=i+1}^M e^{-\frac{|\tilde{\xi}_j^{(k)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right)$$

quantités connues à ce stade de l'algorithme ;

D'autre part :

$$\begin{aligned}
B &= \sum_{l=1}^L \frac{1}{L} \left(\prod_{j=1}^{i-1} e^{-\frac{|\tilde{\xi}_j^{(k+1)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \times \prod_{j=i+1}^M e^{-\frac{|\tilde{\xi}_j^{(k)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right) \\
&= \sum_{l=1}^L \frac{1}{L} \left(\prod_{j=1}^{i-1} e^{-\frac{|\tilde{\xi}_j^{(k+1)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \times \prod_{j=i+1}^M e^{-\frac{|\tilde{\xi}_j^{(k)} - \tilde{\xi}_j^{(l)}|^2}{2h^2}} \right) \\
&= \sum_{l=1}^L \frac{1}{L} \beta_l \\
&= \alpha
\end{aligned} \tag{2.37}$$

en notant $\alpha = \sum_{l=1}^L \frac{1}{L} \beta_l$, une quantité également connue à ce stade.

$$A \stackrel{\text{noté}}{=} \sum_{l=1}^L \frac{1}{L} \beta_l \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi}h} \quad B = \alpha$$

Ainsi :

$$\begin{aligned} p_{Z_i}(x_i) &= \frac{A}{B} \\ &= \frac{\sum_{l=1}^L \frac{1}{L} \beta_l \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi}h}}{\alpha} \\ &= \sum_{l=1}^L \frac{\beta_l}{L\alpha} \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi}h} \\ &= \sum_{l=1}^L \gamma_l \frac{e^{-\frac{|x_i - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi}h} \quad \text{avec } \gamma_l = \frac{\beta_l}{L\alpha} \end{aligned} \quad (2.38)$$

L'expression de cette densité prend une forme similaire à celle de la méthode précédente. Ainsi, en posant :

— V une v.a. sur \mathbb{R} de loi de probabilité discrète $\forall l \in \llbracket 1; L \rrbracket P(V = l) = \gamma_l$.

(en effet : $\sum_{l=1}^L \gamma_l = \sum_{l=1}^L \frac{\beta_l}{L\alpha} = \frac{1}{\alpha} \sum_{l=1}^L \frac{\beta_l}{L} = \frac{1}{\alpha} \alpha = 1$)

— $\forall l \in \llbracket 1; L \rrbracket$, Y_l une v.a. sur \mathbb{R} de densité, $\forall t \in \mathbb{R}$, $p_{Y_l}(t) = \frac{e^{-\frac{|t - \tilde{\xi}_i^{(l)}|^2}{2h^2}}}{\sqrt{2\pi}h}$

Alors, par une démonstration identique à celle de la p. 32, il vient que $Z_i \stackrel{\mathcal{L}}{=} Y_V$ (égalité des variables en loi), et le même algorithme permet de générer Z_i par la simulation de Y_V à partir de V .

Comparaison et synthèse

Pour comparer ces deux méthodes de simulation, examinons à quels résultats elles aboutissent sur un exemple. Choisissons deux variables mutuellement dépendantes⁷, rencontrées à l'occasion des validations du chapitre 3. Des échantillons de 10 000 réalisations de chacune des variables sont simulés suivant l'une et l'autre des deux méthodes proposées. La figure 2.2 compare leurs densités de probabilité respectives (estimées bien sûr grâce à l'estimateur à noyaux).

7. Second et troisième mode de la 5^e classe des 10 000 signaux de réponses présentés au § 3.4.

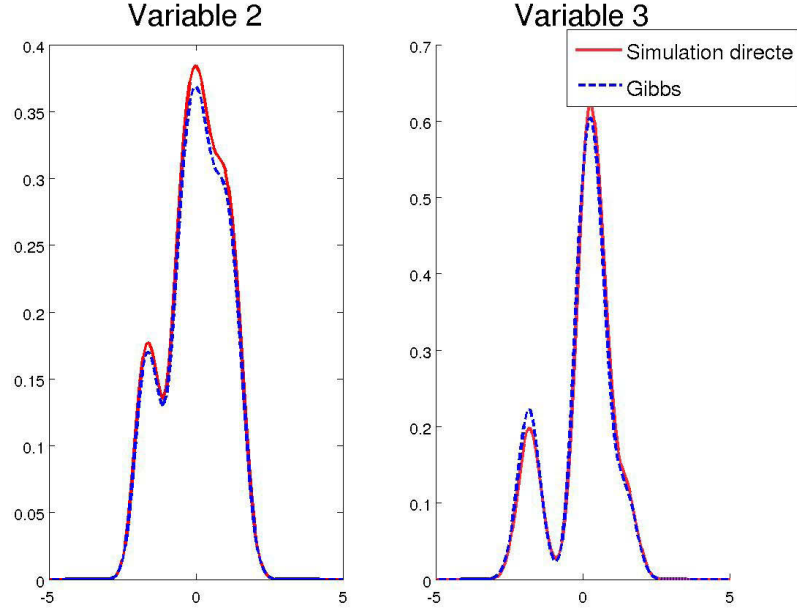


FIGURE 2.2 – Exemples d’estimations, selon les deux méthodes, de la densité de probabilité de variables simulées.

Les résultats sont très similaires, et nous prenons le parti d’utiliser désormais la plus simple des deux méthodes.

Par la suite, les variables ξ_α sont conjointement simulées selon la méthode directe.

2.4 Tri par classes et estimation des courbes de fragilité

Pour construire des courbes de fragilité sans hypothèse de forme *a priori*, et surtout sans utiliser la méthode du *scaling*⁸, nous proposons de regrouper les différents accélérogrammes disponibles par « classes » d’indicateur croissant de nocivité sismique, identifiées par leur valeur moyenne. Ces valeurs moyennes servent d’abscisses à la courbe de fragilité qu’il s’agit de construire.

Après « passage dans la structure », chaque classe de signaux d’entrée correspond *in fine* à une classe de signaux de réponse, éventuellement enrichie comme nous le verrons au § 2.5.3. Ainsi, pour une classe d’intensité α , on note :

- N_α le nombre de signaux de réponse de la classe ;
- n_α^{def} le nombre de signaux de réponse correspondant à une défaillance de la structure.

La probabilité de défaillance recherchée (l’ordonnée de la courbe de fragilité) s’estime alors par :

$$\hat{P}(\alpha) = \frac{n_\alpha^{def}}{N_\alpha} \quad (2.39)$$

8. Voir § 1.3.3

Le regroupement des signaux d'entrée pourrait *a priori* s'effectuer selon des méthodes simples. Les classes d'indicateur pourraient être choisies de largeurs identiques : l'intervalle entre les valeurs minimum et maximum de l'indicateur de nocivité est divisé en un certain nombre de sous-intervalles de mêmes largeurs. Les signaux ainsi regroupés forment les classes, de largeurs identiques mais de tailles de population variables. Certaines classes risquent alors de ne contenir que quelques signaux, ce qui est insuffisant pour mener des calculs de Monte-Carlo ; les classes pourraient également être choisies de tailles de population identiques : dans ce cas, les valeurs moyennes des indicateurs sont souvent mal distribuées, et la courbe de fragilité devient difficile à exploiter.

Pour palier à ces deux problèmes, nous proposons de réaliser le regroupement suivant un algorithme de partitionnement de données, en l'occurrence celui des k -moyennes, présenté dans [32]. Dans l'exemple d'un tri de 10 000 signaux, cette méthode aboutit à 10 classes contenant chacune un minimum de 500 signaux, ce qui est largement exploitable dans un contexte Monte-Carlo. Nous évoquerons ultérieurement la robustesse de ce partage.

Il s'agit de trier en n classes un nombre N_s de signaux, selon la valeur de l'indicateur de nocivité sismique α . Considérons que $\forall i \in \llbracket 1; N_s \rrbracket$, l'indicateur du signal i est α_i : on dispose donc d'un ensemble d'observations $(\alpha_1, \dots, \alpha_{N_s})$, que l'on cherche à organiser selon une partition $S = \{S_1, \dots, S_n\}$, chaque classe S_i affichant une moyenne μ_i . Le regroupement est fait de manière à minimiser la somme des distances « intra-classes » :

$$\sum_{i=1}^n \sum_{\alpha_j \in S_i} |\alpha_j - \mu_i|^2 \quad (2.40)$$

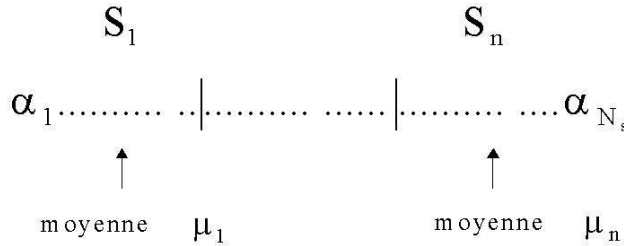


FIGURE 2.3 – Schématisation des observations α_i et de leur partitionnement suivant n classes S_i de moyennes μ_i .

La famille S optimisant cette quantité se détermine en suivant l'algorithme suivant :

1. les moyennes initiales $(\mu_1^{(0)}, \dots, \mu_n^{(0)})$ sont choisies au hasard ;

2. au rang k , on suppose disposer d'une partition $S^{(k)} = \{S_1^{(k)}, \dots, S_n^{(k)}\}$ de moyennes $(\mu_1^{(k)}, \dots, \mu_n^{(k)})$.

Chaque signal est alors « ré-assigné » dans la classe dont le point moyenne lui est le plus proche, **ce qui revient à faire « bouger les cloisons » entre les classes** (cf. figure 2.3). Ainsi :

$$S_i^{(k+1)} = \{\alpha_j : |\alpha_j - \mu_i^{(k)}| \leq |\alpha_j - \mu_l^{(k)}| \text{ pour tout } l \in \llbracket 1; n \rrbracket\} \quad (2.41)$$

La moyenne de chaque classe est ensuite mise à jour :

$$\mu_i^{(k+1)} = \frac{1}{|S_i^{(k+1)}|} \sum_{\alpha_j \in S_i^{(k+1)}} \alpha_j \quad (2.42)$$

3. la seconde étape est répétée jusqu'à « convergence ».

Le nombre de partitions possibles étant fini, et chaque étape aboutissant à une meilleure partition que la précédente, l'algorithme converge donc toujours en temps fini, vers un minimum **local**. Le résultat est donc influencé par le choix des moyennes initiales $(\mu_1^{(0)}, \dots, \mu_n^{(0)})$, qu'il est sage de choisir comme milieux des **classes équiréparties**. Cet algorithme est lent à converger, et sa vitesse n'est pas estimable dans le cas général. En pratique, l'opérateur pose donc lui-même les conditions d'arrêt et réclame la fin de la boucle, soit après un nombre maximum d'itérations, soit lorsque la quantité (2.40) à minimiser passe en-dessous d'un seuil pré-défini, choisi en fonction des ordres de grandeur du problème.

2.5 Prise en compte des incertitudes

Notre méthodologie fait apparaître différents types d'incertitudes, qui doivent être traitées :

- l'incertitude épistémique, due à une certaine ignorance de la structure étudiée ;
- l'incertitude statistique, due aux faible nombre de signaux disponibles ;
- les incertitudes sur le modèle K-L.

2.5.1 Incertitude épistémique

L'incertitude épistémique concentre toute la méconnaissance de la structure. Dans nos applications, nous considérons des modèles mécaniques, dont les paramètres peuvent être modélisés par des variables aléatoires : ils sont donc différents à chaque calcul de structure (voir la mise en œuvre p. 48).

2.5.2 Incertitude statistique

Pour intégrer l'incertitude statistique liée au faible nombre de données disponibles, la courbe de fragilité obtenue avec les seules réponses issues de la famille des signaux d'entrée est accompagnée d'un intervalle de confiance au seuil de confiance 95%. Rappelons que si une probabilité p est estimée par une valeur \hat{p}

après N calculs avec N grand, l'intervalle de confiance au seuil de confiance de 95% s'estime par :

$$\left[\hat{p} - 1,96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{N}} , \quad \hat{p} + 1,96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{N}} \right] \quad (2.43)$$

2.5.3 Incertitudes sur le modèle K-L

Le modèle K-L utilise de nombreuses approximations : troncature de la somme de Karhunen-Loève, estimation d'une matrice d'auto-corrélation, résolution d'un problème aux valeurs propres par une méthode de Galerkin, méthode des rectangles. Nous proposons deux approches pour contenir les incertitudes qui en découlent.

Echantillonnages successifs (*bootstrap*)

Une manière d'intégrer toute l'incertitude sur le modèle consiste à réaliser des échantillonnages successifs sur les signaux, soit d'entrée, soit de réponse, pour construire un intervalle autour de la courbe de la fragilité.

S'il est convenu de les réaliser sur les signaux d'entrée, il suffit d'en tirer au hasard un certain nombre sans remise. Ils sont utilisés pour définir une première courbe de fragilité. L'opération est répétée un certain nombre de fois jusqu'à obtenir autant de courbes différentes. Leurs quantiles fournissent alors un intervalle encadrant la courbe.

Si on souhaite au contraire les réaliser sur les réponses, on en tire de la même manière une sous-famille, au hasard et sans remise. Cette sous-famille est alors elle-même enrichie, c'est-à-dire utilisée comme base pour générer 5 000 nouvelles réponses avec la méthode K-L. Cette opération est répétée un certain nombre de fois jusqu'à obtenir autant de courbes différentes. L'intervalle se construit alors par un calcul de quantiles.

La différence de traitement s'explique par la volonté de limiter le nombre de calculs de structures, et donc celui de signaux d'entrée utilisés. A l'inverse, le nombre de signaux de réponse peut augmenter sans que cela ne prête à conséquence.

Resimulation de la matrice d'autocorrélation

Dans un contexte où les signaux disponibles sont très peu nombreux, l'incertitude viendra de manière principale de l'estimation des matrices d'autocorrélation, notées R de manière générique. Il est possible d'affiner le *bootstrap* précédent de la manière suivante, qui nous est propre.

Dans chacune des classes de réponses, au moment de l'enrichissement, nous commençons par « resimuler » l'estimée de R en suivant la méthode proposée ci-après. C'est cette nouvelle matrice qui est utilisée pour générer les nouveaux signaux, selon la méthode K-L.

Méthode

Soit une classe déterminée de signaux de réponse.

Soit $Y = \{Y(t)\}_{t \in [0, T]}$ le processus stochastique (une fois centré) modélisant le mouvement de la structure soumise à un séisme de la classe en question. Si n réalisations de ce processus $Y = \{Y(t)\}_{t \in [0, T]}$ sont disponibles, ceci fournit n processus indépendants Y^1, \dots, Y^n , réalisés chacun une fois selon y^1, \dots, y^n .

Au besoin, les processus $\{Y^k\}_{k \in \llbracket 1, n \rrbracket}$ seront confondus avec leurs représentations vectorielles $Y^k \in \mathcal{M}_{1, n}(\mathbb{R})$

Pour chaque couple d'instants (t_i, t_j) ⁹, **on cherche à simuler la v.a. de l'estimée de la fonction d'auto-corrélation** :

$$\hat{R}_{i,j} = \frac{1}{n-1} \sum_{k=1}^n Y^k(t_i) Y^k(t_j) \quad (2.44)$$

$\forall k \in \llbracket 1; n \rrbracket$ soit $Z^k = (Y^k)^\top Y^k$

avec donc $\forall (i, j) \in \llbracket 1; N \rrbracket^2$: $Z_{ij}^k = Y^k(t_i) Y^k(t_j)$ de réalisation $z_{ij}^k = y^k(t_i) y^k(t_j)$

Soit alors :

$$\bar{Z}_{ij} = \frac{1}{n} \sum_{k=1}^n Z_{ij}^k \quad \text{et} \quad \bar{z}_{ij} = \frac{1}{n} \sum_{k=1}^n z_{ij}^k \quad (2.45)$$

Donc

$$\boxed{\bar{Z}_{i,j} = \frac{n-1}{n} \hat{R}_{i,j}} \quad (2.46)$$

La matrice $\bar{Z} = \frac{n-1}{n} \hat{R}$ étant symétrique, sa partie supérieure est « déroulée » ligne par ligne, ce qui construit un vecteur aléatoire \bar{U} de longueur $\frac{N(N+1)}{2}$:

$$\bar{U} = (\bar{Z}_{1,1}, \dots, \bar{Z}_{1,N}, \bar{Z}_{2,1}, \dots, \bar{Z}_{2,N}, \bar{Z}_{3,1}, \dots, \bar{Z}_{3,N}, \dots, \dots, \bar{Z}_{N,N}) \quad (2.47)$$

La simulation de \bar{U} revient à celle de tous les $\bar{Z}_{i,j}$, c'est-à-dire de la matrice \hat{R} .

Le vecteur \bar{U} est la moyenne de n vecteurs aléatoires différents U^k , tous de longueur $\frac{N(N+1)}{2}$:

$$U^k = (Z_{1,1}^k, \dots, Z_{1,N}^k, Z_{2,1}^k, \dots, Z_{2,N}^k, Z_{3,1}^k, \dots, Z_{3,N}^k, \dots, \dots, Z_{N,N}^k) \quad (2.48)$$

dont chacun dispose d'une réalisation :

$$u^k = (z_{1,1}^k, \dots, z_{1,N}^k, z_{2,1}^k, \dots, z_{2,N}^k, z_{3,1}^k, \dots, z_{3,N}^k, \dots, \dots, z_{N,N}^k) \quad (2.49)$$

9. Voir les notations sur l'échantillonnage p. 93

En notant alors :

- $\mathbf{m} = (\bar{z}_{1,1}, \dots, \bar{z}_{1,N}, \bar{z}_{2,1}, \dots, \bar{z}_{2,N}, \bar{z}_{3,1}, \dots, \bar{z}_{3,N}, \dots, \dots, \bar{z}_{N,N})$
- $\Gamma = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{u}^k - \mathbf{m})^T (\mathbf{u}^k - \mathbf{m})$

et en supposant que \mathbf{m} et Γ sont respectivement le vecteur moyenne commun et la matrice d'autocorrélation commune des vecteurs \mathbf{U}^k , le théorème central limite¹⁰ assure alors que :

$$\frac{\bar{\mathbf{U}} - \mathbf{m}}{1/\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}, \Gamma) \quad (2.50)$$

ce qui invite alors à simuler $\bar{\mathbf{U}}$ comme une gaussienne multidimensionnelle $\mathcal{N}\left(\mathbf{m}, \frac{1}{n}\Gamma\right)$.

Comme cela a déjà été précisé, la simulation de $\bar{\mathbf{U}}$ revient à celle des $\bar{Z}_{i,j}$, et donc de \hat{R} .

Méthode « simplifiée »

La méthode précédente est difficilement réalisable en pratique. Dans le cadre des petites expérimentations que nous présenterons au chapitre 3, le vecteur $\bar{\mathbf{U}}$ est de longueur $\frac{N(N+1)}{2}$, fournissant pour $N = 2000$ une matrice de variance-covariance de plus de 4.10^{12} coefficients ! ce qui est inenvisageable avec Matlab. Nous proposons ici une « seconde méthode » pour contourner ce problème, sous réserve, évidemment, de quelques hypothèses simplificatrices.

L'expression $\hat{R}_{i,j} = \frac{1}{n-1} \sum_{k=1}^n Z_{ij}^k$, valable pour tout couple (i,j) , est résumée par l'égalité matricielle :

$$\hat{R} = \frac{1}{n-1} \sum_{k=1}^n \mathbf{Z}^k \quad (2.51)$$

Cette matrice est symétrique réelle, et donc diagonalisable dans une base orthonormée. Ainsi, il existe \mathbf{P} une matrice orthogonale et \mathbf{D} une matrice diagonale vérifiant :

$$\hat{R} = \mathbf{P} \mathbf{D} \mathbf{P}^T \quad (2.52)$$

L'hypothèse est faite que **la matrice \mathbf{P} est déterministe**, *i.e.* qu'elle est la même pour toutes les réalisations de $\hat{R} : \mathbf{P} \in O_N(\mathbb{R})$. Seule \mathbf{D} reste une matrice aléatoire, que nous allons chercher à simuler, ce qui reviendra

10. Voir p. 89

évidemment à simuler \hat{R} . Ainsi :

$$\begin{aligned}
D &= P^\top \hat{R} P \\
&= P^\top \left(\frac{1}{n-1} \sum_{k=1}^n Z^k \right) P \\
&= \frac{1}{n-1} \sum_{k=1}^n P^\top Z^k P \\
&= \frac{n}{n-1} \times \frac{1}{n} \sum_{k=1}^n P^\top Z^k P
\end{aligned} \tag{2.53}$$

On note à présent $\forall k \in \llbracket 1; n \rrbracket$:

$$\begin{aligned}
W^k &= P^\top Z^k P \\
&= P^\top (Y^k)^\top Y^k P \\
&= (Y^k P)^\top Y^k P
\end{aligned} \tag{2.54}$$

Donc $\forall i \in \llbracket 1; N \rrbracket$

$$\begin{aligned}
D_{ii} &= \frac{n}{n-1} \times \frac{1}{n} \sum_{k=1}^n \left[(Y^k P)^\top Y^k P \right]_{ii} \\
&= \frac{n}{n-1} \times \frac{1}{n} \sum_{k=1}^n [Y^k P]_i^2
\end{aligned} \tag{2.55}$$

Soit alors le vecteur aléatoire \bar{V} de longueur N :

$$\bar{V} = \left(\frac{1}{n} \sum_{k=1}^n [Y^k P]_1^2, \dots, \frac{1}{n} \sum_{k=1}^n [Y^k P]_N^2 \right) \tag{2.56}$$

qui est en fait la moyenne de n vecteurs aléatoires différents V^k , tous de longueur N :

$$V^k = \left([Y^k P]_1^2, \dots, [Y^k P]_N^2 \right) \tag{2.57}$$

dont chacun dispose d'une réalisation :

$$v^k = \left([y^k P]_1^2, \dots, [y^k P]_N^2 \right) \tag{2.58}$$

En notant alors :

- $m = \left(\frac{1}{n} \sum_{k=1}^n [y^k P]_1^2, \dots, \frac{1}{n} \sum_{k=1}^n [y^k P]_N^2 \right)$
- $\Gamma = \frac{1}{n-1} \sum_{k=1}^n (v^k - m)^T (v^k - m)$

et en supposant que \mathbf{m} et Γ sont respectivement le vecteur moyenne commun et la matrice d'autocorrélation commune des vecteurs \mathbf{V}^k , le théorème central limite assure là-aussi que :

$$\frac{\bar{\mathbf{V}} - \mathbf{m}}{1/\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}, \Gamma) \quad (2.59)$$

ce qui invite donc ici aussi à simuler $\bar{\mathbf{V}}$ comme une gaussienne multidimensionnelle $\mathcal{N}\left(\mathbf{m}, \frac{1}{n}\Gamma\right)$.

Avoir simulé $\bar{\mathbf{V}}$ revient à avoir simulé tous les D_{ij} , c'est-à-dire la matrice \hat{R} .

Commentaires

Une fois la matrice R resimulée, elle est utilisée pour générer une nouvelle famille de variables ξ_α , selon l'équation (2.12) du § 2.1.2 :

$$\xi_\alpha^{(l)} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^N x_\Gamma^{(l)}(t_j) \bar{\phi}_\alpha(t_j) \Delta t$$

Mais on constate que cette simulation exige d'utiliser les signaux $x_\Gamma^{(l)}$, qui seront donc toujours les mêmes à chaque resimulation de la matrice R . Il est légitime de s'attendre d'ores et déjà à un dysfonctionnement lors de la mise en pratique de cette méthode.

Chapitre 3

Analyse et validation de la méthodologie sur un cas simple

Le programme de validation de la méthodologie proposée au chapitre 2 s'articule selon quatre étapes distinctes. Il examine tour à tour :

- la représentativité du modèle K-L pour la génération de signaux d'entrée. Des accélérogrammes artificiels sont simulés à partir d'une base de données, et leur concordance est examinée ;
- la robustesse du tri par classes, en en faisant varier le nombre et la taille ;
- la représentativité du modèle K-L pour la génération de signaux de réponse. Deux structures simples sont soumises à des signaux d'entrée, regroupés par classes. Les calculs mécaniques fournissent alors des classes de signaux de réponse, chacune utilisée elle-même comme base pour générer une seconde fois des réponses. Ceci permet de vérifier leur concordance ;
- la prise en compte des différentes incertitudes (voir §. 2.5 et 3.5.).

Rappelons que lors d'études industrielles, par souci de temps de calcul, seuls quelques centaines de signaux d'excitation peuvent en pratique être soumis au modèle numérique de la structure. Un grand nombre de signaux sont toutefois simulés dans ce chapitre, à des fins de validation.

Tous nos calculs sont réalisés avec le logiciel Matlab.

3.1 Données

3.1.1 Base de signaux réels

La base d'accélérogrammes **réels** utilisée ici est un sous-ensemble de la base de signaux européens ESD (*European Strong-Motion Database* [19]). Les signaux retenus pour notre approche sont sensés constituer

des réalisations d'un même processus stochastique $\{\Gamma(t)\}_{t \in [0, T]}$ qui modélise l'accélération du sol pour le scénario sismique¹ retenu. On considère le cas suivant :

- distance à la source $0 < \Delta < 20$ km.
- magnitude $5.5 < M < 6.5$;

97 accélérographes adéquats ont été retenus, après quelques légères modifications : la valeur en $t = 0$ est fixée à 0 ; la suite est tronquée pour qu'au premier instant corresponde la première valeur dépassant, en valeur absolue, une valeur seuil, fixée arbitrairement à $5.10^{-4}g = 4.9 \times 10^{-3} \text{m.s}^{-2}$.

De plus, enregistrés avec un pas de temps $\Delta t = 10^{-2}$ s insuffisant pour les calculs mécaniques, les signaux subissent un « ré-échantillonnage² » selon un nouveau pas de temps $\Delta t' = \frac{dt}{9} = 1.1.10^{-3}$ s.

3.1.2 Modèles mécaniques

Deux modèles mécaniques simples [20] ont été retenus pour notre programme de validation :

- l'oscillateur linéaire à un degré de liberté, de pulsation propre ω_0 et de taux d'amortissement ξ_0 . Lorsqu'il est soumis à l'accélération $\{\gamma(t)\}_{t \in [0, T]}$, sa réponse en déplacement $\{y(t)\}_{t \in [0, T]}$ vérifie :

$$\ddot{y}(t) + 2\omega_0\xi_0\dot{y}(t) + \omega_0^2y(t) = -\gamma(t) \quad (3.1)$$

- l'oscillateur non-linéaire dans lequel la réponse, également notée $\{y(t)\}_{t \in [0, T]}$, à une accélération $\{\gamma(t)\}_{t \in [0, T]}$, vérifie :

$$\ddot{y}(t) + 2\xi_0\omega_0\dot{y}(t) + F(y(t), \alpha) = -\gamma(t) \quad (3.2)$$

La fonction F est représentée sur figure 3.1. Les différents paramètres du modèle sont la pulsation ω_0 , le taux d'amortissement réduit ξ_0 , le rapport des pentes α , la limite d'élasticité en déplacement y_{nl} (notée Y sur le schéma de la figure 3.1). Notons que $\omega_0 = \sqrt{\frac{k}{m}}$ avec une masse $m = 1$.

1. Voir p. 15.

2. Voir annexe B

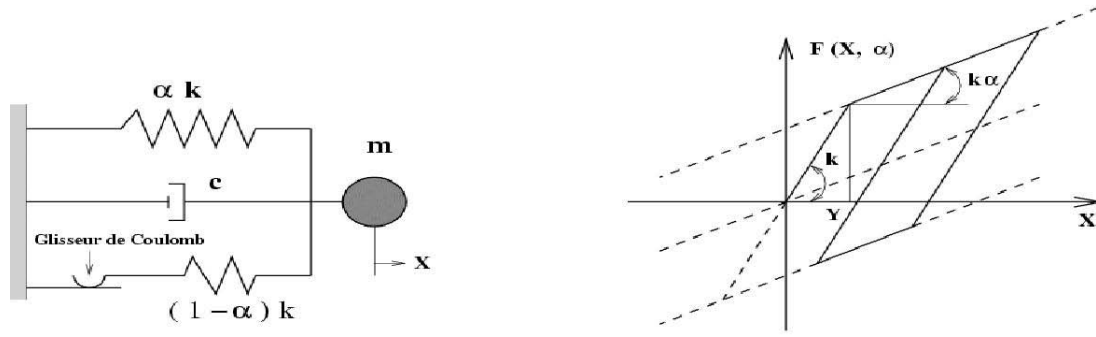


FIGURE 3.1 – *Modèle rhéologique de l'oscillateur avec écrouissage cinématique.*

Les valeurs des paramètres sont fixées aux alentours des ordres de grandeur représentatifs d'une structure en acier :

- la pulsation propre $\omega_0 = 2\pi 5 \text{ rad.s}^{-1}$, qui correspond à une fréquence propre $f_0 = 5 \text{ Hz}$;
- l'amortissement $\xi_0 = 5\%$;
- le rapport des pentes $\alpha = 0.2$;
- la vitesse et le déplacement initiaux $\dot{y}(0)$ et $y(0)$, choisis nuls.

Limite d'élasticité et critère de défaillance

Le choix de la limite d'élasticité en déplacement y_{nl} passe par une étude de dimensionnement. 10 000 accélérogrammes simulés grâce à la méthode K-L³ sont regroupés selon 10 classes de PGA, et leurs SRO moyens à 5% sont illustrés sur la figure 3.10 de la p.56. La valeur de y_{nl} est choisie de telle sorte que la structure n'adopte un comportement non-linéaire qu'à partir des signaux de la troisième classe de PGA : ainsi, sur le SRO moyen de la troisième classe, on note y_{nl} la valeur vérifiant $\omega_0^2 y_{nl} = SRO(\omega_0)$, soit :

$$y_{nl} = 3.10^{-3} \text{ m} \quad (3.3)$$

Soit $\{Y(t)\}_{t \in [0, T]}$ le processus stochastique représentant la réponse en déplacement de la structure soumise au processus $\{\Gamma(t)\}_{t \in [0, T]}$. La probabilité (conditionnelle) de la défaillance de la structure en fonction de l'indicateur de nocivité du séisme auquel elle est soumise s'écrit, avec b le « critère de défaillance », ou « déplacement maximum admissible », défini au §1.3.1 :

$$P_f(\alpha) = P\left(\max_{t \in [0, T]} |Y(t)| > b \mid \text{intensité } \alpha\right) \quad (3.4)$$

3. Il s'agit de la famille E_1 présentée au § 3.2

La **ductilité** admissible est alors définie par $\mu = \frac{b}{y_{nl}}$. Le critère b est ordinairement choisi pour que $\mu = 2$, d'où $b = 2y_{nl}$, soit :

$$b = 6.10^{-3}m \quad (3.5)$$

Ce critère de défaillance est utilisé pour les cas linéaire et non-linéaire.

Prise en compte des incertitudes sur les paramètres de la structure

L'incertitude sur les paramètres ω_0 , ξ_0 , α et y_{nl} , peut être prise en compte en les modélisant par des variables aléatoires, comme mentionné au § 2.5. Si un paramètre était choisi dans un premier temps à une valeur μ , il est alors modélisé comme une v.a. uniforme $\mathcal{U}([90\%\mu, 110\%\mu])$.

Cette modélisation est utilisée dans tout ce qui suit, sans que cela ne soit rappelé.

3.2 Représentativité pour la simulation de signaux d'entrée

La première vocation du modèle K-L dans cette thèse est de permettre l'enrichissement d'une famille d'accélérogrammes, qu'ils soient réels ou synthétiques. Trois familles de signaux d'entrée sont utilisées :

- la base de « signaux réels » E_0 , constituée des 97 accélérogrammes présentés au § 3.1.1 ;
- une famille E_1 de 10 000 « signaux simulés », générés à partir de E_0 selon la méthode K-L ;
- une famille E_2 de 10 000 « signaux resimulés », générés à partir des 10 000 signaux simulés de la famille E_1 . Cette troisième famille contribuera à vérifier l'efficacité de la méthode lorsqu'un grand nombre de signaux de départ est disponible.

3.2.1 Paramètres statistiques globaux

Assurons-nous en premier lieu que ces trois familles d'accélérogrammes réels, simulés et resimulés présentent les mêmes caractéristiques, à savoir :

- leur dispersion, évaluée à l'aide de l'estimateur de la variance ;
- les indicateurs de nocivité du § 1.1 et les Spectres de réponse.

Variance en fonction du temps

Si un processus $\{\Gamma(t)\}_{t \in [0, T]}$ est réalisé L fois selon $\{\gamma^{(l)}\}_{l \in \llbracket 1; L \rrbracket}$, sa moyenne $\mu_\Gamma(t) = E[\Gamma(t)]$ s'estime sans biais sur les valeurs $\{t_k = (k - 1)\Delta t\}_{k \in \llbracket 1; N \rrbracket}$ grâce à :

$$\tilde{\mu}_\Gamma(t_k) = \frac{1}{L} \sum_{l=1}^L \gamma^{(l)}(t_k) \quad (3.6)$$

Sa variance $V_{\Gamma}(t) = E[(\Gamma(t) - \mu_{\Gamma}(t))^2]$ s'estime, également sans biais, par :

$$\tilde{V}_{\Gamma}(t_k) = \frac{1}{L-1} \sum_{l=1}^L (\gamma^{(l)}(t_k) - \tilde{\mu}_{\Gamma}(t_k))^2 \quad (3.7)$$

La figure 3.2 compare alors les variances (estimées) des signaux réels de E_0 et des signaux simulés de E_1 . De la même façon, la figure 3.3 compare les variances (estimées) des signaux simulés de E_1 et des signaux resimulés de E_2 . Pour plus de lisibilité, seul l'intervalle de temps $[0,10]$ est représenté. Dans les deux cas, la concordance est très bonne.

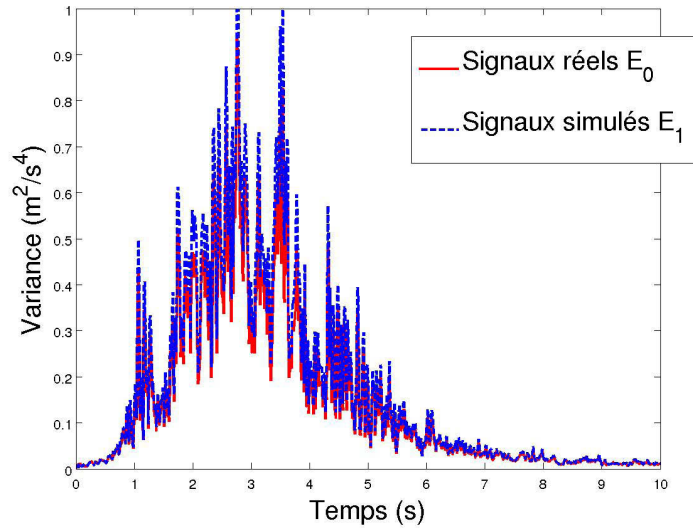


FIGURE 3.2 – *Variance estimée d'accélérogrammes réels et simulés, sur l'intervalle de temps $[0s; 10s]$.*

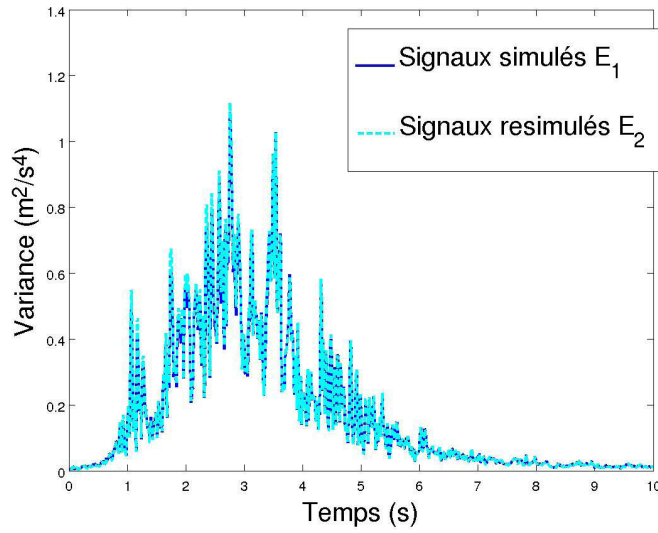


FIGURE 3.3 – *Variance estimée d'accélérogrammes simulés et resimulés, sur l'intervalle de temps [0s; 10s].*

Paramètres de nocivité

La « nocivité sismique » d'un accélérogramme peut s'évaluer à l'aide des indicateurs définis au paragraphe 1.1 : le PGA, le CAV, l'IA, le T_{sm} et la PSA pour un amortissement et à une fréquence donnés, par exemple 0.05 et 5 Hz, valeurs en adéquation avec les ordres de grandeurs des fréquences des structures d'intérêt.

Les tableaux 3.1, 3.2 et 3.3 présentent ces paramètres en termes de moyenne, de médiane et d'écart-type, pour les trois familles d'accélérogrammes réels, simulés et resimulés.

	PGA ($m.s^{-2}$)	CAV ($m.s^{-1}$)	Arias ($m.s^{-1}$)	PSA ($m.s^{-2}$)	$T_{sm}(s)$
Moyenne	1,42	2,12	0,15	3,31	6,45
Médiane	1	1,65	0,05	2,41	6,31
Ecart-type	1,21	1,87	0,32	3,17	2,63

TABLE 3.1 – *Paramètres statistiques de la famille E_0 des 97 accélérogrammes réels.*

	PGA ($m.s^{-2}$)	CAV ($m.s^{-1}$)	Arias ($m.s^{-1}$)	PSA ($m.s^{-2}$)	$T_{sm}(s)$
Moyenne	1,65	2,57	0,18	3,89	5,89
Médiane	1,23	2,04	0,07	2,88	5,8
Ecart-type	1,16	1,71	0,33	2,96	1,39

TABLE 3.2 – *Paramètres statistiques de la famille E_1 des 10 000 accélérogrammes simulés.*

	PGA ($m.s^{-2}$)	CAV ($m.s^{-1}$)	Arias ($m.s^{-1}$)	PSA ($m.s^{-2}$)	$T_{sm}(s)$
Moyenne	1,68	2,59	0,18	3,97	5,74
Médiane	1,26	2,05	0,07	2,94	5,65
Ecart-type	1,16	1,71	0,33	2,98	1,3

TABLE 3.3 – Paramètres statistiques de la famille E_2 des 10 000 accélérogrammes resimulés.

Une légère différence s’observe entre l’échantillon des signaux réels E_0 et celui des signaux synthétiques E_1 , surtout pour l’indicateur T_{sm} , problème déjà souligné par les articles [89] et [52]. Le problème vient sans doute du fait que la corrélation $R(t,t)$ est mal représentée. La représentativité est très bonne de E_1 à E_2 , situation dans laquelle la famille de signaux de départ était cette fois très fournie.

Nous privilégierons pour la construction des courbes de fragilité l’indicateur PGA, le plus utilisé dans l’ingénierie parasismique.

La figure 3.4 compare alors, pour un amortissement de 5%, les SRO⁴ moyens, médians et quantiles à 15 et 85% de la famille des signaux réels E_0 , simulés E_1 et resimulés E_2 .

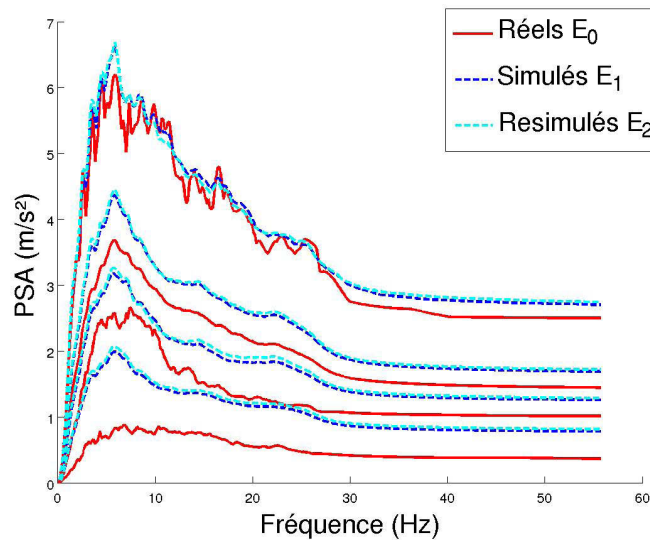


FIGURE 3.4 – Quantiles de SRO moyens à 5%, issus des données réelles E_0 , simulées E_1 et resimulées E_2 : respectivement de haut en bas les quantiles à 85%, les moyennes, les médianes et les quantiles à 15%.

Les conclusions sont identiques : si la concordance n’est pas totale entre E_0 et E_1 , elle l’est entre E_1 et E_2 . Ceci incite à penser que le modèle K-L serait très opérationnel dans une situation où le nombre de signaux disponibles serait relativement élevé.

4. Cf. § 1.1

3.2.2 Modes propres de Karhunen-Loève

Comparons à présent les différentes « fonctions propres » (au sens de Karhunen-Loève) de la matrice d'autocorrélation R empirique des signaux réels, simulés et resimulés. La figure 3.5 compare les premières valeurs propres obtenues pour les trois familles de signaux.

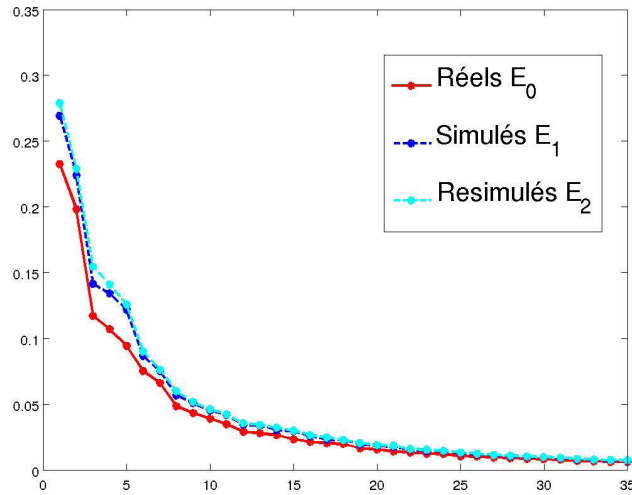


FIGURE 3.5 – Valeurs propres du modèle K - L obtenues pour les données réelles, simulées et resimulées.

Comme précédemment, le passage de E_0 et E_1 est acceptable et très bon de E_1 à E_2 . Comparons à présent les fonctions propres des trois familles de signaux. Dans un souci de clarté, la figure 3.6 ne présente que les 12 premières fonctions propres, et seulement sur l'intervalle de temps $[9s, 10s]$, au moment où la variance du processus est la plus forte. La concordance est très bonne.

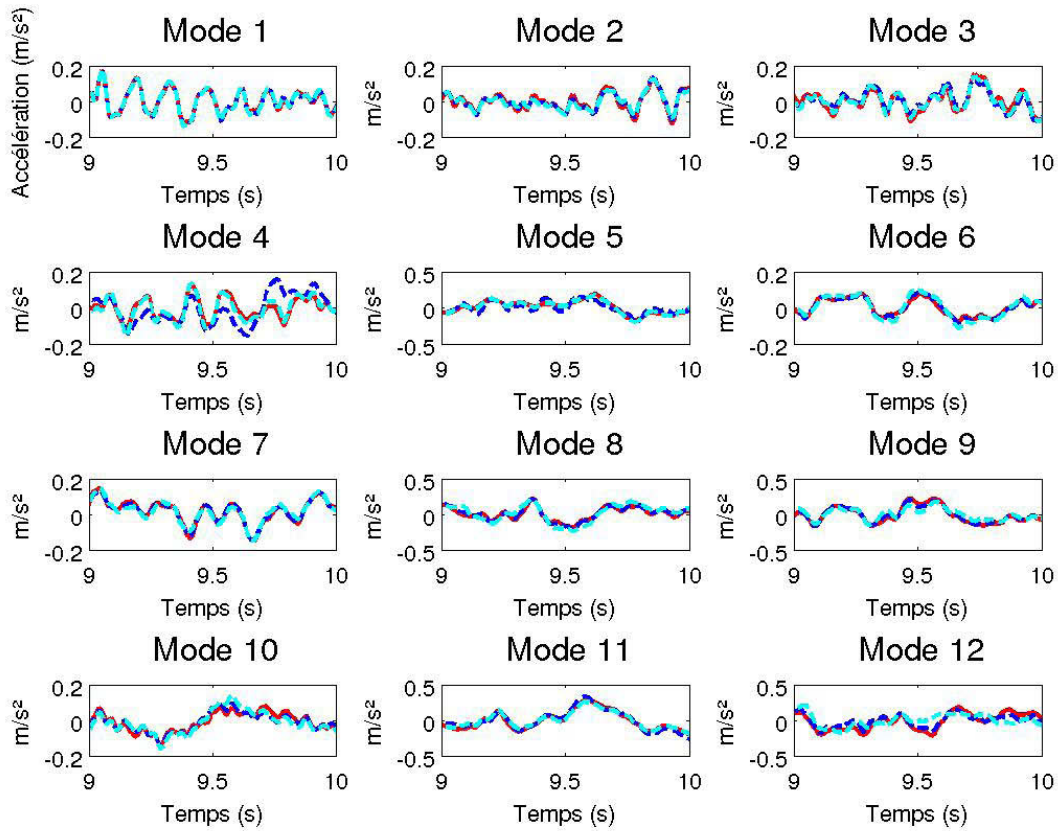


FIGURE 3.6 – Premières fonctions propres pour les données réelles, simulées et resimulés (respectivement en rouge, bleu et cyan).

3.2.3 Taille de base intermédiaire

Nous ne disposons que de 97 signaux réels, mais avons remarqué que les résultats seraient très bons si la famille de signaux de départ était mieux fournie. Mais pourrait-on améliorer les résultats avec une base de taille intermédiaire, supérieure à 100 mais inférieure à 10 000 ? Examinons ce qui se passerait si la base de départ était composée de, par exemple, 500 signaux. On considère deux familles :

- T_1 : **500 signaux simulés**, générés à partir de E_0 . Ses courbes sont affichées en **bleu** ;
- T_2 : **10 000 signaux resimulés**, générés à partir de T_1 . Ses courbes sont affichées en **magenta**.

Examinons les mêmes indicateurs que précédemment :

- les paramètres statistiques aux tableaux 3.4 et 3.5 ;
- les quantiles des SRO figure 3.7 ;
- les valeurs propres figure 3.8 ;
- les premiers modes propres figure 3.9.

	PGA ($m.s^{-2}$)	CAV ($m.s^{-1}$)	Arias ($m.s^{-1}$)	PSA ($m.s^{-2}$)	$T_{sm}(s)$
Moyenne	1,66	2,59	0,18	3,93	5,91
Médiane	1,18	2,01	0,07	2,83	5,8
Ecart-type	1,17	1,8	0,35	3,11	1,41

TABLE 3.4 – Paramètres statistiques de la famille T_1 des 500 accélérogrammes simulés.

	PGA ($m.s^{-2}$)	CAV ($m.s^{-1}$)	Arias ($m.s^{-1}$)	PSA ($m.s^{-2}$)	$T_{sm}(s)$
Moyenne	1,74	2,73	0,19	4,13	5,78
Médiane	1,32	2,19	0,08	3,08	5,63
Ecart-type	1,12	1,71	0,33	2,99	1,25

TABLE 3.5 – Paramètres statistiques de la famille T_2 des 10 000 accélérogrammes resimulés.

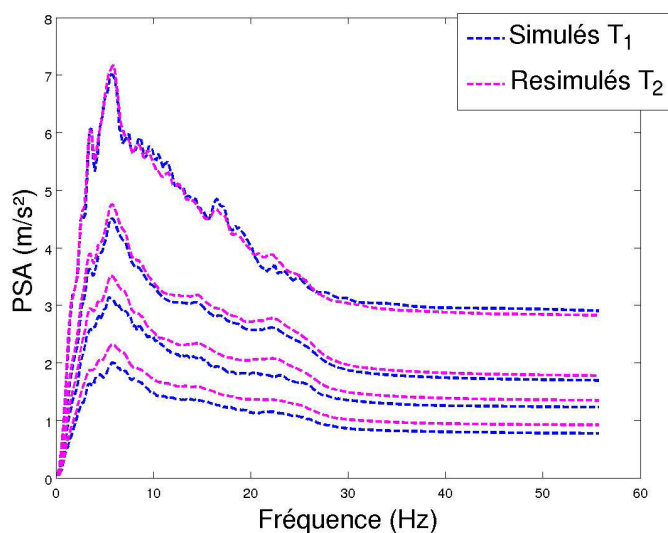


FIGURE 3.7 – Quantiles de SRO moyens à 5%, issus des données simulées T_1 et resimulées T_2 : respectivement de haut en bas les quantiles à 85%, les moyennes, les médianes et les quantiles à 15%.

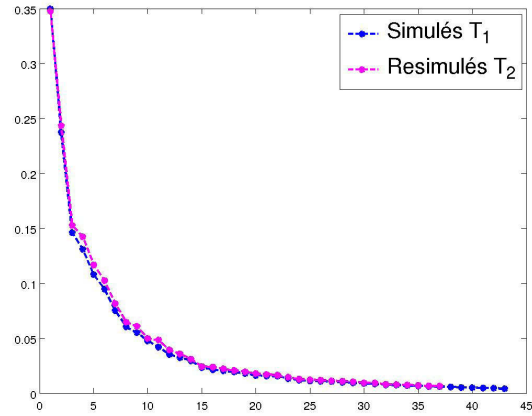


FIGURE 3.8 – Valeurs propres du modèle K-L obtenues pour les données simulées T_1 et resimulées T_2 .

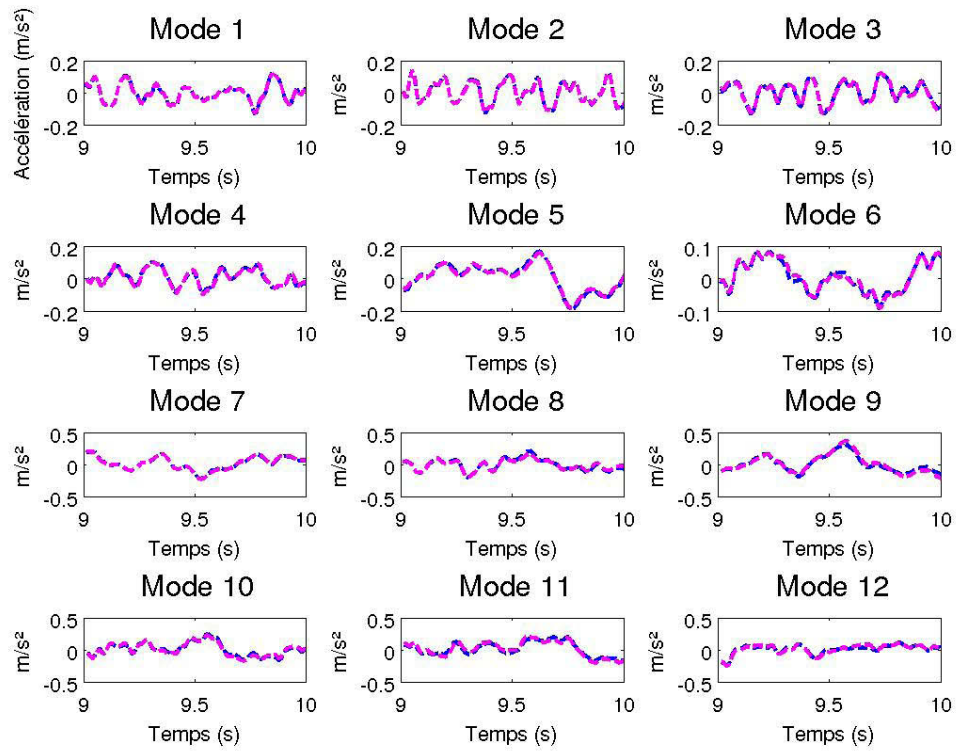


FIGURE 3.9 – Premières fonctions propres pour les données simulées T_1 et resimulées T_2 (respectivement en bleu et magenta).

La concordance est bonne dans tous les cas : les résultats s'avèrent satisfaisants dès que la famille de signaux de base possèdent quelques centaines de signaux. Nous poursuivons toutefois notre étude avec les 10 000 signaux de la famille E_1 , dans un but de validation.

3.3 Tri par classes

En vue de construire des courbes de fragilité, les signaux d'entrée sont regroupés par « classes » selon la méthode des k -moyennes présentée au paragraphe 2.4. La démarche est identique quel que soit le paramètre statistique utilisé. Dans ce qui suit, les signaux simulés sont donc comme nous l'avons déjà précisé triés suivant leur valeur de PGA, l'indicateur le plus utilisé en génie parasismique.

Les 10 000 signaux de la famille E_1 sont regroupés selon 10 classes de PGA, identifiées par leur PGA moyen. La figure 3.10 présente les SRO à 5% moyens de chaque classe de PGA.

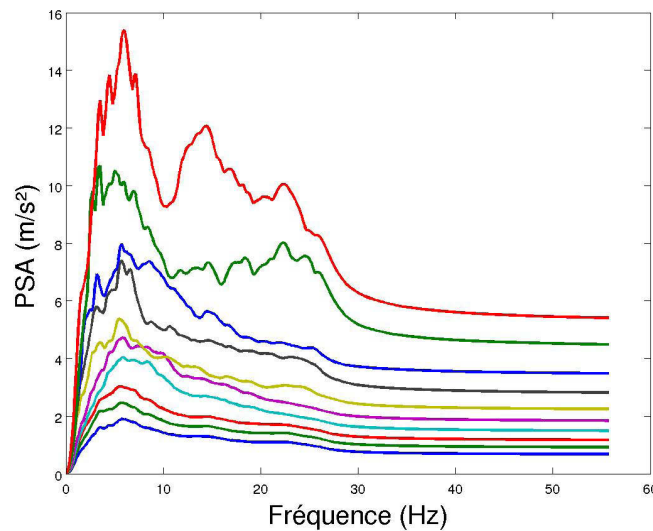


FIGURE 3.10 – Spectres de Réponse d'Oscillateur (à 5%) moyens des 10 000 signaux simulés pour 10 classes croissantes de PGA.

Le contenu spectral dépend de l'intensité de la classe (ici en PGA). C'est cette dépendance qui nous avait incité à écarter la méthode du *scaling* : celle-ci suppose un comportement identique de la réponse linéaire avec l'augmentation de l'intensité sismique des signaux d'entrée.

3.3.1 Corrélation intra-classes des signaux d'entrée

Un problème de taille apparaît à ce stade. Parmi les signaux simulés de la famille E_1 , certains semblent se ressembler très fortement. Ceci conduit en pratique à réaliser plusieurs fois le même « calcul de structure », et

donc à créer de l'information redondante et inutile. Pour une première illustration nous décidons, dans chacune des deux classes 9 et 10, d'isoler les signaux présentant un fort coefficient de corrélation, par exemple supérieur à 0,9, **avec le premier mode** (prépondérant) de la classe. La figure 3.11 présente les résultats sur l'intervalle de temps [4; 5] pour plus de lisibilité.

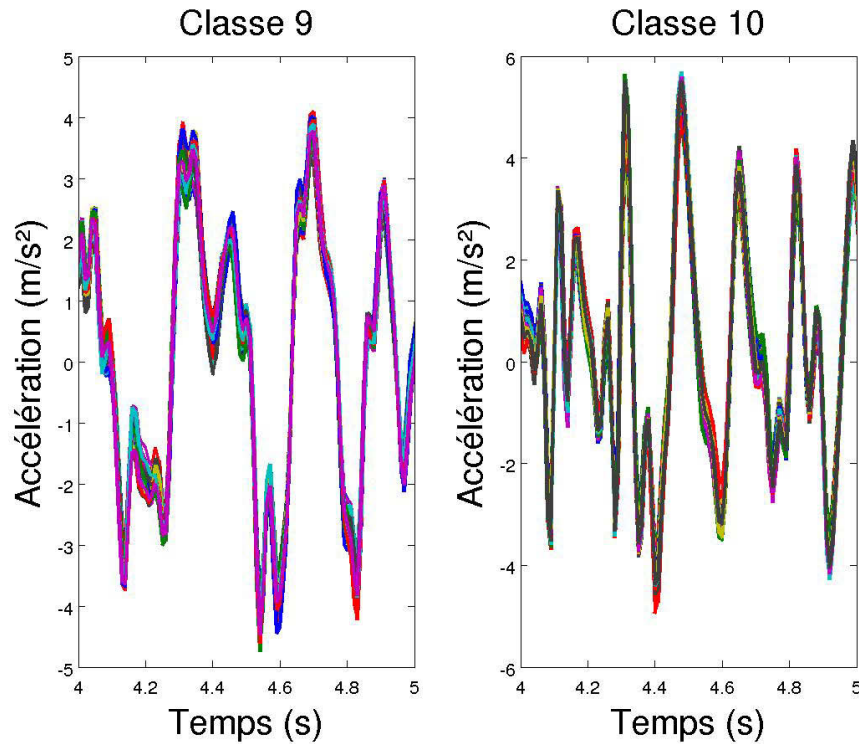


FIGURE 3.11 – Exemples de signaux simulés ressemblants : 54 de la classe n°9 (à gauche) et 98 de la classe n°10 (à droite).

- la classe 9 comporte 54 signaux très ressemblants, soit près de 25% de ses 218 signaux ;
- la classe 10 comporte 98 signaux très ressemblants, soit plus de 20% de ses 464 signaux.

Le tableau 3.6 présente alors, dans chaque classe, les signaux présentant un coefficient de corrélation supérieur à 0,9 avec **un des cinq premiers modes**.

Classe	1	2	3	4	5	6	7	8	9	10
PGA moyen (m.s^{-2})	0,67	0,9	1,15	1,46	1,81	2,21	2,76	3,42	4,37	5,27
Effectif	1652	2127	1523	1255	1039	730	528	464	218	464
Signaux « ressemblants »	0	0	0	0	25	76	83	45	159	100
Pourcentage	0%	0%	0%	0%	2,4%	10,4%	15,7%	9,7%	73%	21,6%

TABLE 3.6 – Effectif et pourcentage, dans chaque classe, de signaux d'entrée « ressemblants » (corrélation supérieure à 0,9 avec un des cinq premiers modes).

A l'examen de ce tableau, le problème semble apparaître pour les classes comportant peu de signaux d'origine, qui sont en fait les classes de fort PGA. Nous y reviendrons dans la partie 3.6 de ce chapitre.

3.3.2 Détermination de la taille des classes

Pour ce qui suit, nous souhaitons regrouper les signaux d'entrée selon différents nombres de classes, jusqu'à 100, et les 10 000 signaux de la famille E_1 ne pourront fournir des classes suffisamment fournies pour des résultats exploitables. **La base de signaux réels E_0 est alors utilisée pour générer, non plus 10 000, mais exceptionnellement 100 000 accélérogrammes artificiels**, successivement regroupés en 5, 10, 20, 30, 50 puis 100 classes de PGA.

Pour ces différents partages en classes, les figures 3.12 et 3.13 présentent, sans intervalle de confiance, les courbes de fragilité issues des structures respectivement linéaire et non-linéaire soumises à ces 100 000 signaux.

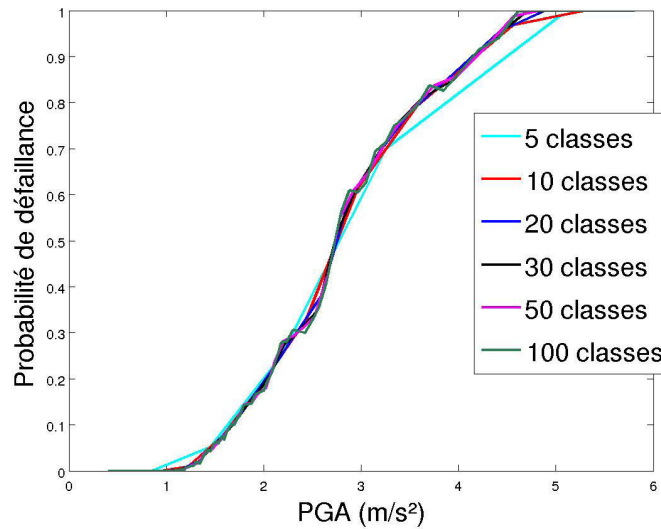


FIGURE 3.12 – Structure *linéaire* : courbes de fragilité obtenues pour différents partages en classes de PGA.

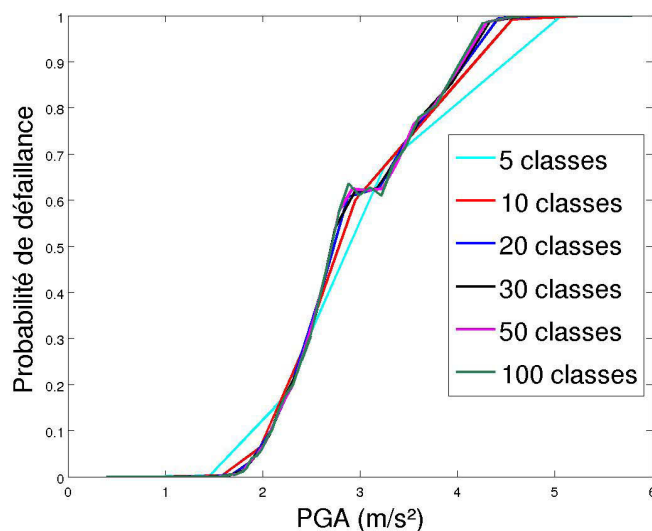


FIGURE 3.13 – Structure **non-linéaire** : courbes de fragilité obtenues pour différents partages en classes de PGA.

Dans les deux situations, il est raisonnable de considérer que le résultat est inchangé selon le nombre de classes, à partir d'au moins 10 classes. Dans la suite de cette validation, nous réaliserons un regroupement en 10 classes de PGA.

3.4 Représentativité pour la simulation de signaux de réponse

Enrichir la base de signaux d'entrée augmente d'autant plus le nombre de calculs mécaniques, coûteux en temps. Pour en limiter le nombre, nous proposons à présent d'enrichir plutôt la famille des signaux de réponse de chaque classe. Mais si le modèle K-L semble convenir pour enrichir une famille d'accélérogrammes, des signaux filtrés pourraient présenter un contenu spectral particulier et doivent être étudiés séparément.

Les 10 000 signaux d'entrée synthétiques de la famille E_1 ont été triés selon 10 classes de PGA. Nous disposons donc, pour chacune des deux structures, de 10 familles de signaux de réponse, appelées « réponses d'origine » par la suite. Nous commençons par examiner si, oui ou non, la « ressemblance » par classes des signaux d'entrée, observée au § 3.3.1 s'est propagée aux signaux de réponse, dans le cas linéaire et non-linéaire. Une fois ce point éclairci, chacune des classes sera à son tour enrichie pour former de nouvelles classes de 5 000 signaux de sortie : les résultats seront notés « réponses resimulées » sur les légendes.

3.4.1 Corrélation intra-classes des signaux de réponse

Le tableau 3.7 dresse, pour les cas linéaire et non-linéaire, les mêmes données que le tableau 3.6 : dans chaque classe, il signale les signaux de réponse présentant un coefficient de corrélation supérieur à 0,9 avec un des cinq premiers modes.

Classe	1	2	3	4	5	6	7	8	9	10
PGA moyen (m.s^{-2})	0,67	0,9	1,15	1,46	1,81	2,21	2,76	3,42	4,37	5,27
Effectif	1652	2127	1523	1255	1039	730	528	464	218	464
Rép. linéaires res.	1	1	1	0	0	0	28	0	11	47
Pourcentage (linéaire)	0,1%	0%	0,1%	0%	0%	0%	5,3%	0%	5%	10,1%
Rép. non-linéaires res.	0	0	0	0	0	10	66	0	48	102
Pourcentage (non-linéaire)	0%	0%	0%	0%	0%	1,4%	12,5%	0%	22%	22%

TABLE 3.7 – Effectif et pourcentage, dans chaque classe, de signaux de réponse **linéaire** et **non-linéaire** « ressemblants » (corrélation supérieur à 0,9 avec un des cinq premiers modes).

La prise en compte de l'incertitude sur les paramètres de la structure semble avoir eu partiellement raison des ressemblances observées au sein des classes au § 3.3.1.

3.4.2 Paramètres statistiques globaux

Considérons à titre d'exemple la classe n°5 des réponses à la structure **non-linéaire**. Les variances des deux familles de réponse sont bien reproduites, comme le montre la figure 3.14.

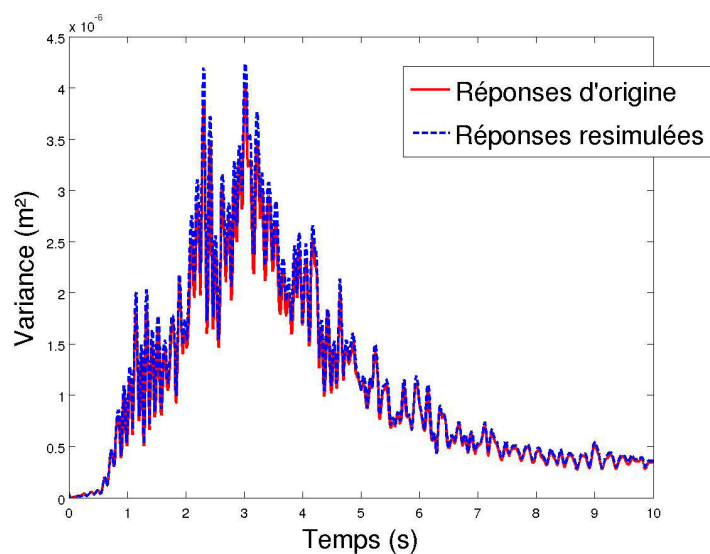


FIGURE 3.14 – Variance estimée du processus de sortie : réponses d'origine et resimulées.

La figure 3.15 assure de même la bonne adéquation de la répartition des maxima des réponses.

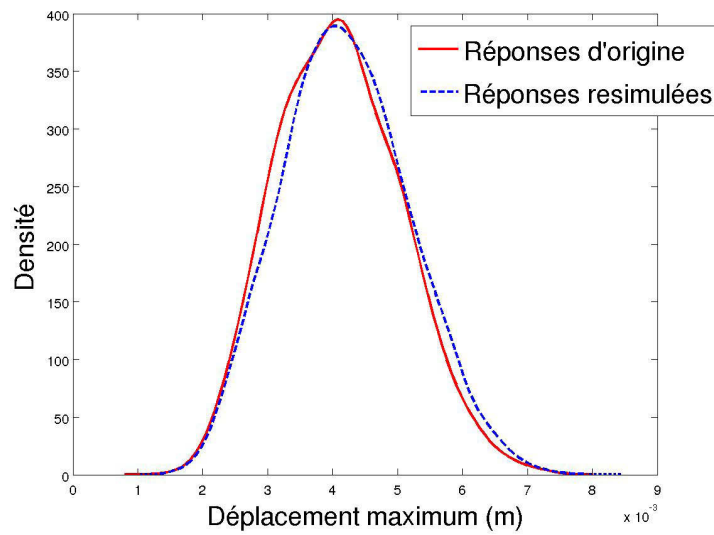


FIGURE 3.15 – Répartition des maxima des réponses d'origine et des réponses resimulées.

La figure 3.16 compare enfin les répartitions des réponses d'origine et resimulées à un instant fixé, celui correspondant au maximum de la variance, à savoir $t = 3,02$ s. La concordance est bonne.

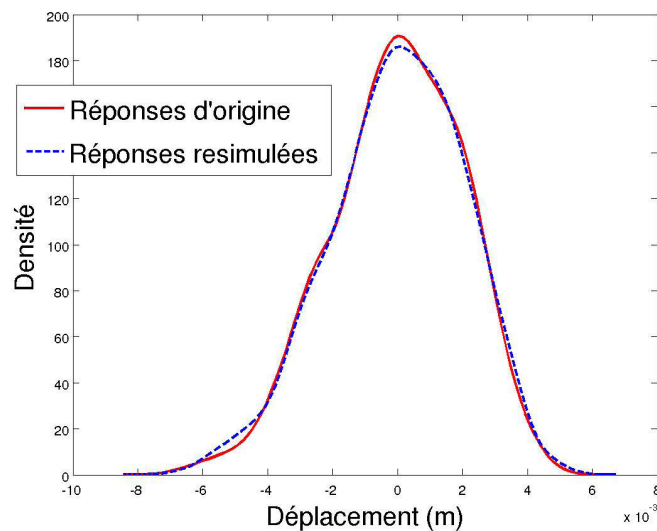


FIGURE 3.16 – Répartition, à l'instant de variance maximum $t = 3,02$ s, des réponses d'origine et resimulées.

3.4.3 Courbes de fragilité

Construisons à présent les courbes de fragilité à partir de ces différentes familles de réponse.

Rappelons que la famille E_1 des 10 000 signaux d'entrée simulés est regroupée selon 10 classes de PGA,

pour une moyenne d'environ un millier de signaux par classe. Les 10 familles de signaux de réponse qui leur correspondent fournissent immédiatement la « **courbe-cible** », qui est représentée **en rouge** dans tout ce qui suit.

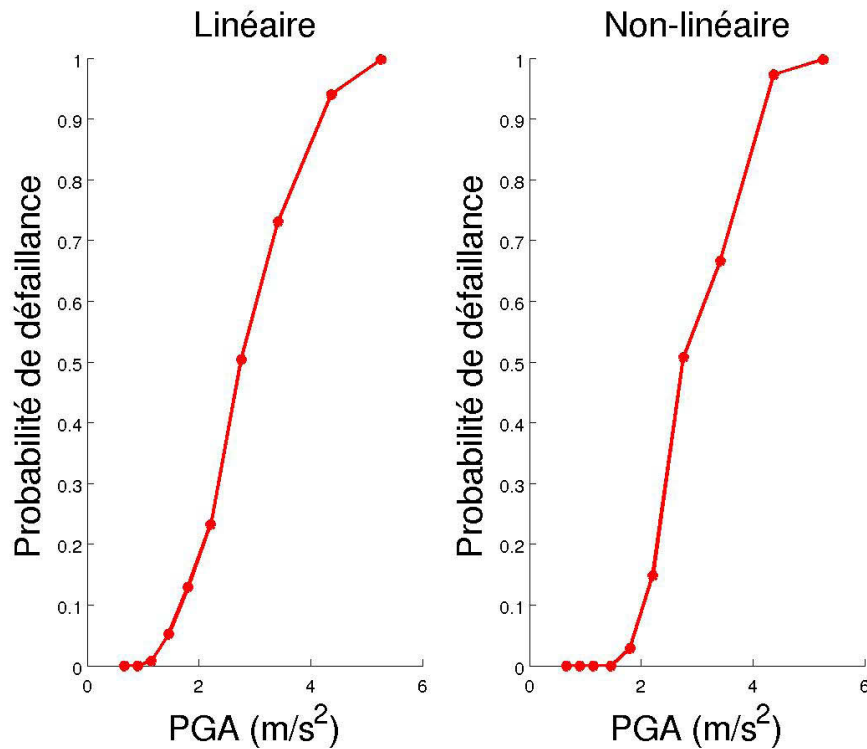


FIGURE 3.17 – « Courbes-cibles », pour le cas linéaire (à gauche) et non-linéaire (à droite).

Ajustement log-normal

Nous examinons sur les figures 3.18 et 3.19 l'adéquation de nos courbes de fragilité avec un modèle log-normal. Déterminons pour cela, par la méthode des moindres carrés, des courbes log-normales s'ajustant au mieux à celles obtenues par simulation de Monte-Carlo.

Les paramètres recueillis dans chacun des cas sont :

- médiane $A_m = 1,01$ et écart-type logarithmique $\beta = 0,33$ dans le cas linéaire ;
- médiane $A_m = 1,06$ et écart-type logarithmique $\beta = 0,28$ dans le cas non-linéaire.

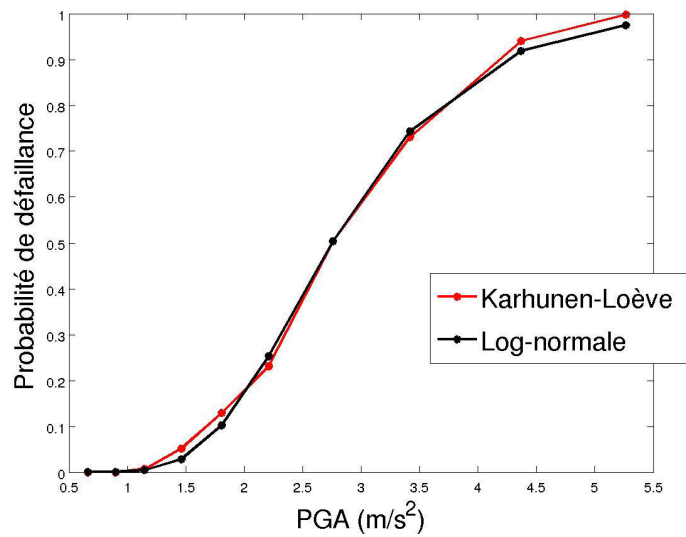


FIGURE 3.18 – Courbe de fragilité, pour l'oscillateur **linéaire**, selon la classe de PGA, et son ajustement log-normal.

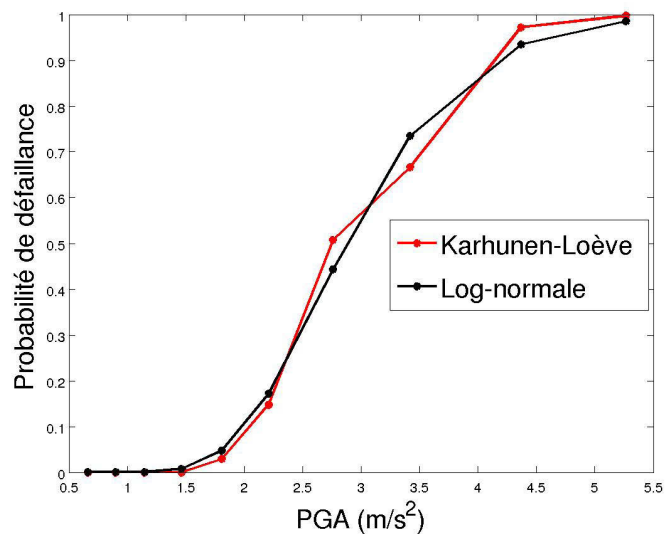


FIGURE 3.19 – Courbe de fragilité, pour l'oscillateur **non-linéaire**, selon la classe de PGA, et son ajustement log-normal.

Dans le cas linéaire, la courbe log-normale est très proche de la courbe rouge obtenue, point par point, par simulation de Monte-Carlo. Mais ce n'est pas le cas pour l'oscillateur non-linéaire. Ceci nous confirme l'intérêt de construire des courbes de fragilité sans l'hypothèse log-normale.

Influence de la prise en compte de la dépendance des coefficients du développement de K-L

Les figures 3.20 et 3.21 présentent :

- en rouge, la courbe-cible attendue ;
- en bleu, la courbe obtenue avec les données issues de la resimulation des réponses : chaque famille de réponses est enrichie de 5 000 nouvelles réponses (voir p. 59) ;
- en noir, les résultats issus de la resimulation mais **sans** prendre en compte la dépendance des coefficients du développement de Karhunen-Loève.

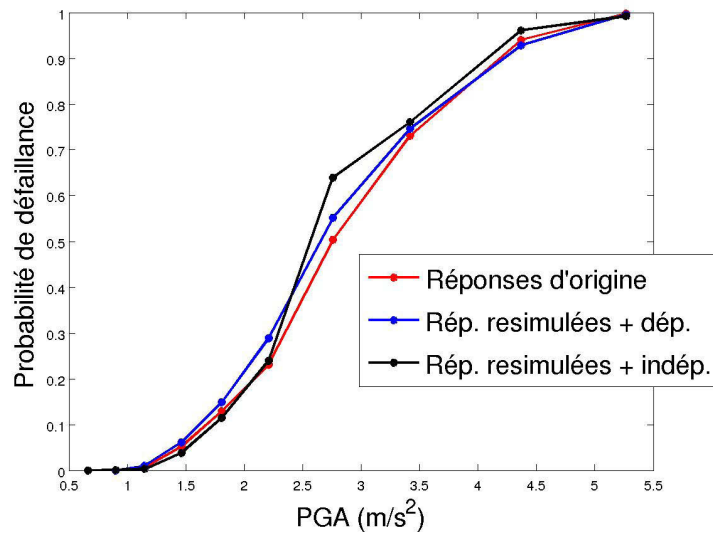


FIGURE 3.20 – Comparaison de l'influence de la dépendance, pour l'oscillateur **linéaire**.

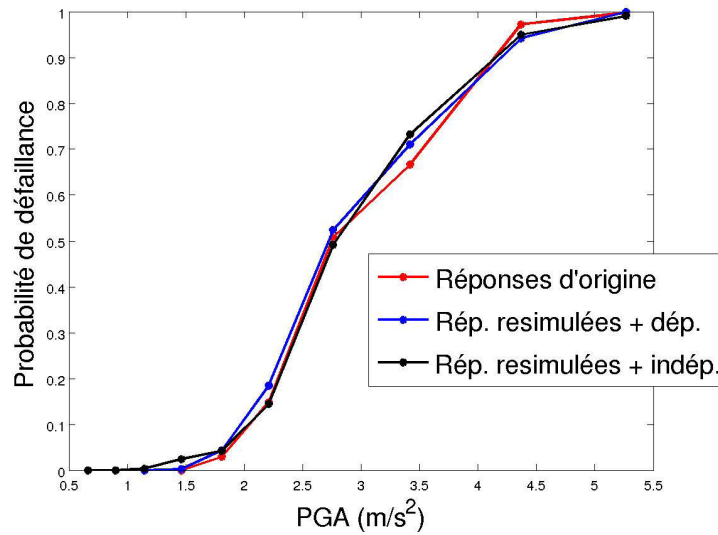


FIGURE 3.21 – Comparaison de l'influence de la dépendance, pour l'oscillateur *non-linéaire*.

Si les deux résultats sont presque identiques dans le cas non-linéaire, celui considérant la dépendance est meilleur dans le cas linéaire, ce qui confirme l'intérêt de prendre en compte la dépendance des coefficients.

3.5 Prise en compte de l'incertitude

Les cas industriels ne peuvent être traités en pratique qu'avec un nombre très réduit de signaux de réponse, situation pour laquelle il convient également de valider notre méthode. Comme précisé au paragraphe 2.5, on souhaite construire autour des courbes des intervalles représentant les différents types d'incertitudes :

- incertitude statistique, liée à la taille des échantillons ;
- incertitude sur le modèle K-L.

3.5.1 Incertitude statistique

Comme précisé p. 38, les courbes-cibles sont accompagnées de l'intervalle de confiance habituel au seuil de confiance 95%, présenté sur les figures 3.22 et 3.23

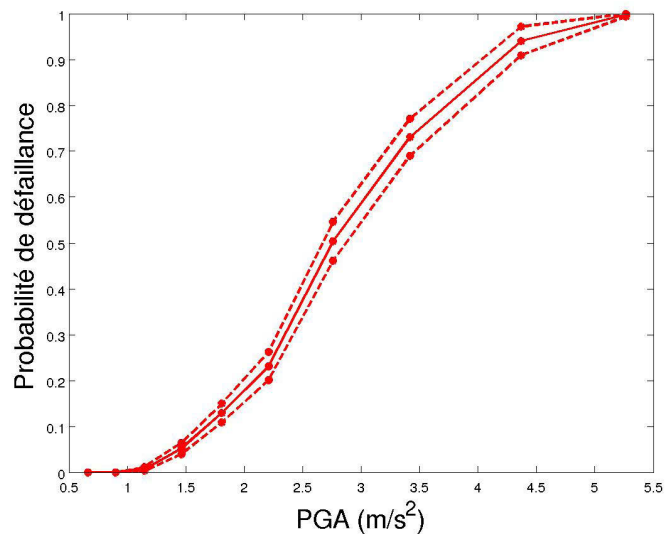


FIGURE 3.22 – Courbe de fragilité, pour l'oscillateur **linéaire**, selon la classe de PGA, et son intervalle de confiance.

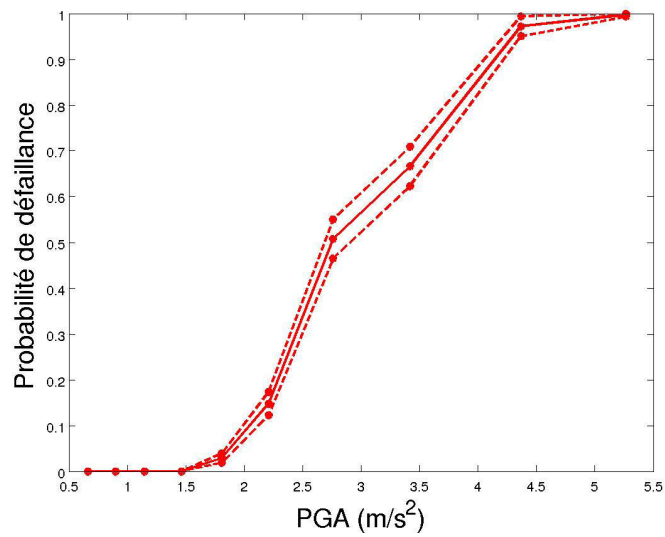


FIGURE 3.23 – Courbe de fragilité, pour l'oscillateur **non-linéaire**, selon la classe de PGA, et son intervalle de confiance.

3.5.2 Incertitude sur le modèle K-L

Resimulation successive des signaux de réponse

Comme précisé p 39, une première possibilité consiste, dans chacune des classes de réponse, à tirer **aléatoirement** une sous-famille de signaux de réponse, utilisée comme base pour générer 5 000 nouvelles réponses avec la méthode K-L. Cette opération est répétée 20 fois, jusqu'à obtenir 20 courbes différentes : leurs quantiles fournissent un intervalle.

Cette opération est réalisée lorsque la sous-famille sélectionnée contient :

- **toutes** les réponses de la classe : « resimulation intégrale » ;
- seulement **une partie** des réponses de la classe : « resimulation sélective ».

Les figures 3.24 et 3.25 comparent :

- en rouge : la courbe-cible et son intervalle de confiance statistique ;
- en bleu : la courbe précédente issue de la resimulation « intégrale » des réponses ;
- en noir (respectivement en magenta) : quantiles à 5% et 95% de 20 courbes de fragilité différentes obtenues en resimulant, dans chaque classe, 5000 réponses, après avoir identifié un modèle KL avec 50 (respectivement 200) réponses de la classe tirées au hasard (sans remise).

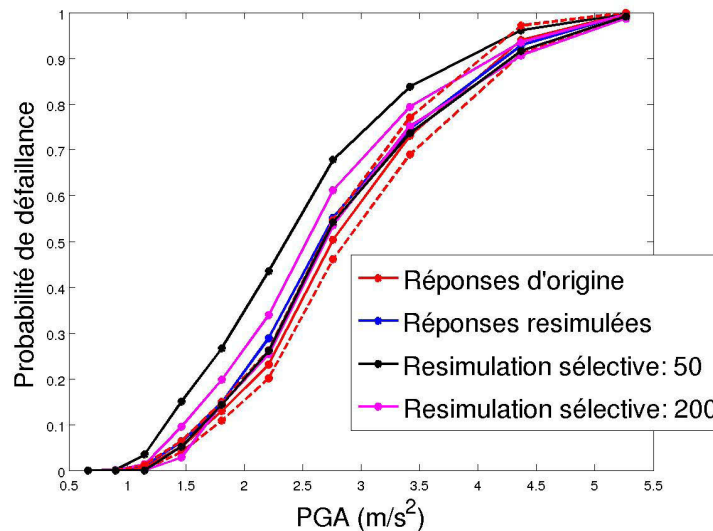


FIGURE 3.24 – Resimulation « intégrale » ou « sélective » des signaux de réponse dans le cas **linéaire**.

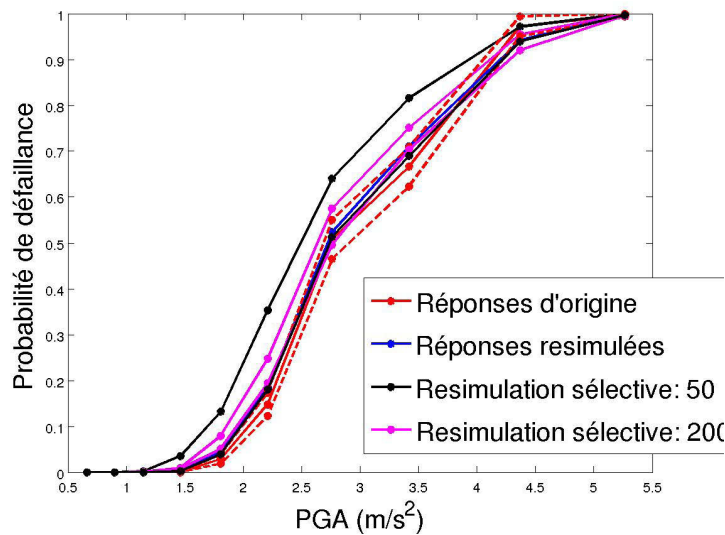


FIGURE 3.25 – *Resimulation « intégrale » ou « sélective » des signaux de réponse dans le cas **non-linéaire**.*

Les résultats sont très encourageants, même si les courbes simulées se placent **au-dessus** de la courbe-cible. Ceci leur confère un caractère **conservatif**, du moins en dehors des fortes classes.

Resimulation de la matrice d'autocorrélation

La méthode précédente ne peut pourtant être réellement mise en pratique, puisqu'il n'y a, dans la réalité, aucun signal à choisir ! Lorsque peu de signaux de réponse sont disponibles, l'intervalle autour de la courbe peut se construire en s'appuyant sur le paragraphe 2.5.3 selon la méthodologie suivante. Dans chacune des classes de signaux de réponse :

- au vue d'un certain nombre de signaux de réponse, la matrice d'autocorrélation est estimée puis « resimulée » selon la procédure que nous avons détaillée ;
- la « nouvelle » matrice d'autocorrélation est utilisée pour générer 5000 signaux de réponse synthétiques avec la méthode K-L ;
- l'opération est renouvelée 20 fois, pour pouvoir au final disposer de 20 courbes de fragilité différentes ; les quantiles associés à 5% et 95% permettront d'estimer un intervalle de confiance.

A titre de vérification, la figure 3.26 présente, dans le cas **non-linéaire**, ce qui arrive lorsque **tous** les signaux de chaque classe participent à l'estimation de la matrice :

- en rouge et bleu, les mêmes courbes que précédemment ;
- en noir : l'intervalle de confiance construit lorsque **l'intégralité** des signaux de chaque classe est utilisée.

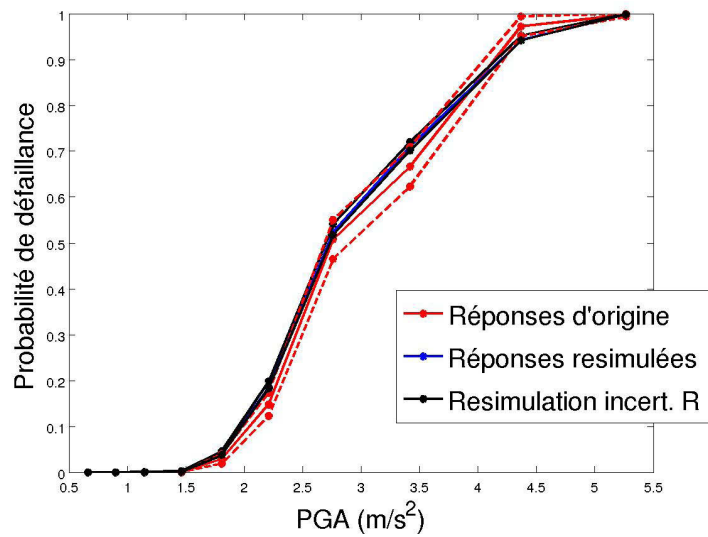


FIGURE 3.26 – *Prise en compte de l'incertitude sur la matrice R dans le cas non-linéaire : utilisation de la **totalité** des signaux de chaque classe de réponse.*

La bonne concordance des résultats invite à présenter la figure 3.27, où **seuls 30 signaux** sont prélevés au hasard dans chaque classe de réponse pour estimer les matrices d'autocorrélation.

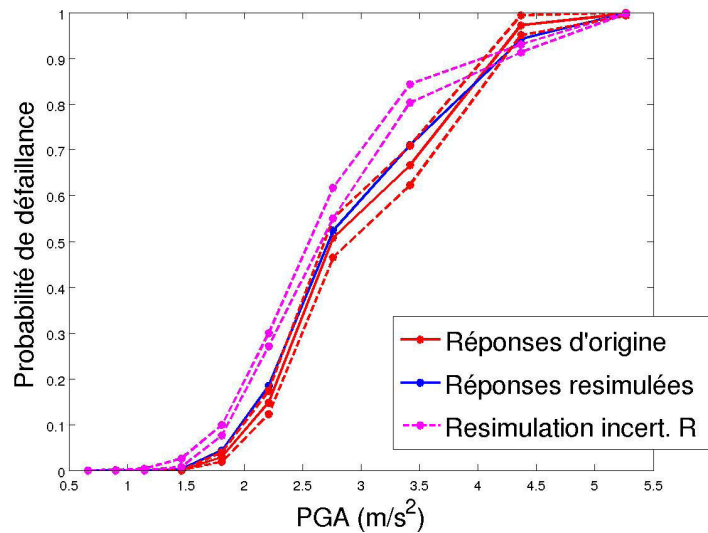


FIGURE 3.27 – *Prise en compte de l'incertitude sur la matrice R dans le cas non-linéaire : utilisation de **seulement 30 signaux** par classe de réponse.*

Ce dernier résultat n'est pas satisfaisant, puisque l'intervalle magenta obtenu n'encadre pas la courbe bleue. Ce fort décalage provient sans doute du problème que nous avons souligné p. 39 : à chaque resimu-

lation de la matrice R , les variables ξ_α sont calculées toujours à partir des mêmes signaux de réponse. Cet aspect pourrait être examiné plus en détail dans un travail ultérieur.

Tirage successif des signaux d'entrée

Nous tentons alors d'inclure toutes les incertitudes de manière globale dans un plus grand intervalle, par le *bootstrap* proposé p. 39. Son application aboutit à la figure 3.28.

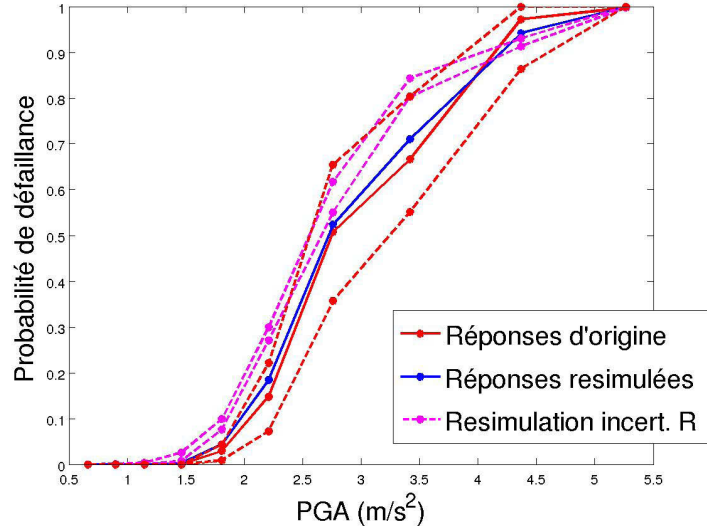


FIGURE 3.28 – Courbes précédentes, et nouvel intervalle (pointillés rouges) issu des tirages successifs des signaux d'entrée.

Cet intervalle ne vaut que pour information, en raison du grand nombre de calculs de structures qu'il réclame. Nous ne l'utiliserons pas au chapitre 4 traitant du cas industriel.

3.6 Discussion

Il est possible de resimuler des accélérogrammes à partir d'une famille de signaux réels grâce à la représentation de Karhunen-Loève. Les nouveaux signaux sont similaires aux anciens en termes de nocivité sismique, contenu spectral, et modes propres prépondérants. Certains signaux simulés se ressemblent toutefois beaucoup, en particulier dans les fortes classes. La probabilité d'occurrence de ces forts séismes ($PGA > 3,5 \text{ m.s}^{-2}$) est toutefois faible, du moins en zone de sismicité moyenne comme la France, puisque le risque sismique est centré sur les faibles séismes. L'introduction de ce mémoire présentait en effet la notion d'« aléa sismique » : celle-ci précise la probabilité de dépassement d'une valeur caractéristique de l'intensité sismique au cours d'une période de référence. Elle décroît, en échelle logarithmique, de manière presque linéaire selon l'intensité sismique (voir par exemple [51]). Cette probabilité est donc en réalité inversement

proportionnelle à une puissance positive de l'intensité sismique et, en France, devient très faible pour les séismes des classes fortes.

Le problème s'est de plus atténué au niveau des signaux de réponse, ceux-ci s'obtenant en prenant compte de l'incertitude sur les paramètres de la structure, comme nous l'avons précisé p. 48. Nos constatations invitent donc à ne pas s'inquiéter des problèmes pouvant survenir au niveau des fortes classes, quitte à ne représenter les courbes que pour les classes faibles, voire intermédiaires.

Après s'être assuré de la robustesse du tri par classes, nous avons pu nous assurer, à l'aide de deux structures simples, qu'il est possible de resimuler les signaux de réponse, sous réserve de prendre en compte la dépendance des coefficients de Karhunen-Loève. Ceci nous permet de construire ces courbes avec moins de signaux d'entrée, proposition intéressante pour des cas industriels où les modèles numériques sont très complexes et les calculs de structures numériquement très coûteux. Dans le but, enfin, de diminuer une nouvelle fois le nombre de signaux d'entrée strictement nécessaire à notre travail, nous avons examiné plusieurs manières d'insérer les incertitudes qui pèsent sur notre modèle. L'une d'elle consiste à manipuler la matrice R qui joue un rôle central dans notre méthode, même si les résultats ne sont pas totalement satisfaisants et mériteraient d'être menés plus avant.

Chapitre 4

Application à un cas réel

Malgré les difficultés soulevées dans la section précédente, nous décidons de tester la méthode sur une situation concrète issue d'un contexte industriel, pour laquelle peu de calculs de structures sont envisageables.

Nous nous dotons pour cela d'un modèle à éléments finis d'une structure réelle. Dans un premier temps, à des fins de validation, les 10 000 signaux d'entrée déjà utilisés au chapitre 3 sont soumis à cette structure. Ceci fournit une première courbe de fragilité, qualifiée de « courbe-cible », accompagnée de son intervalle de confiance statistique. Les calculs de structures nécessaires sont longs, mais indispensables dans le cadre de ce travail, pour permettre d'obtenir cette courbe de référence à laquelle sera comparée toutes les suivantes. Ensuite, comme dans la partie validation de ce mémoire, de nouveaux signaux de réponses sont simulés et la nouvelle courbe de fragilité obtenue est comparée à la courbe-cible. Enfin, l'incertitude de modèle est prise en compte grâce aux considérations sur la matrice R , en vue de construire des intervalles autour de nos courbes.

4.1 Présentation de la ligne

La structure proposée est la ligne de tuyauterie ASG¹, présentée sur la figure 4.1. Elle est composée d'une lyre inclinée dans le plan XOY de 22 degrés par rapport à la verticale Y . Les trois points 12, 19 et 35 matérialisés ici correspondent à des points « pilotes » dans le comportement global de la ligne.

1. Alimentation de Secours en eau du Générateur de Vapeur d'un Réacteur à Eau Pressurisée 1300.

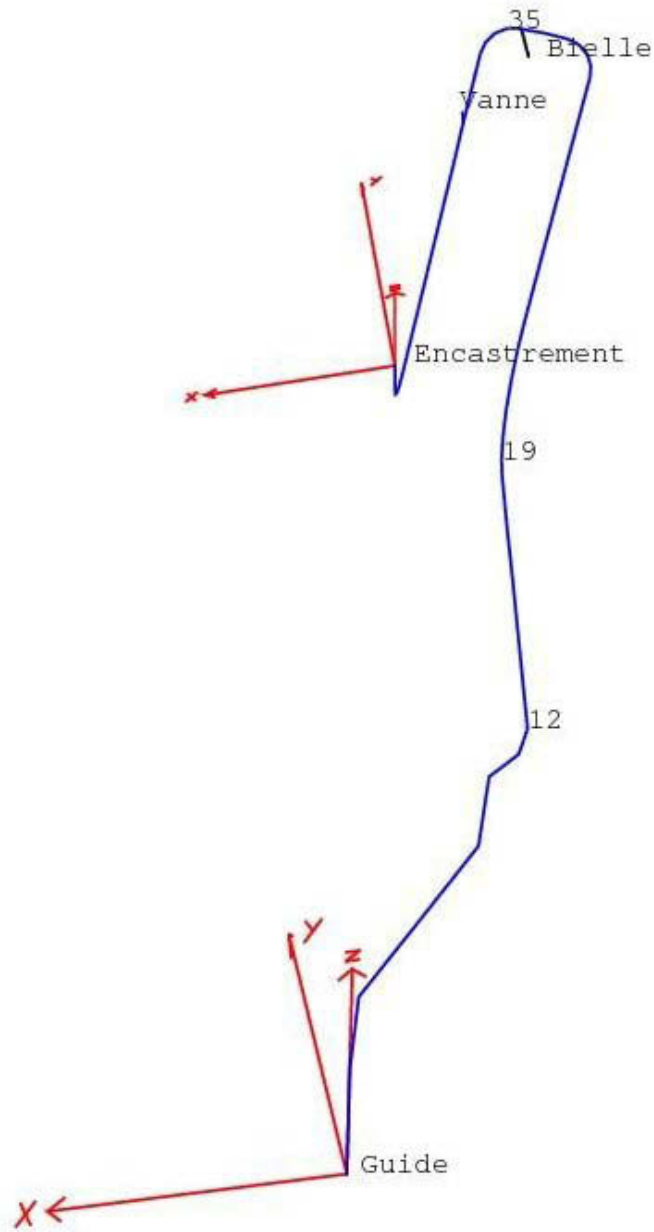


FIGURE 4.1 – *Maillage de la ligne ASG.*

Cette structure est représentée par un modèle à éléments finis réalisé sous Cast3M² à partir de l'élément « Tuyau ». Ses caractéristiques ont été recalées à partir d'essais dynamiques durant lesquels la table vibrante Azalée du CEA-Saclay a été soumise à une sollicitation sismique dans la direction X : les résultats de ces essais, consignés dans le document [7], sont rappelés ici.

2. Logiciel de simulation numérique développé par le Département de Modélisation des Systèmes et Structures du CEA

Caractéristiques

Les principales caractéristiques de la ligne sont indiquées au tableau 4.1.

Longueur suivant X (m)	1,45
Longueur suivant Y (m)	2,46
Longueur suivant Z (m)	4,84
Diamètre extérieur (mm)	114,3
Epaisseur (mm)	8,56
Rayon de courbure des coudes (mm)	152
Rayon de courbures des cintrages (mm)	571
Coefficient de Poisson	0,3
Masse volumique équivalente (kg.m^{-3})	$1,041.10^4$
Module de Young (GPa)	192,36

TABLE 4.1 – Principales caractéristiques de la ligne ASG

La ligne possède trois points de supportage, modélisés selon le tableau 4.2 :

- un encastrement ;
- une bielle avec rotules à chaque extrémité pour bloquer le déplacement selon Z ;
- un guide longitudinal visant à bloquer les déplacements en X et Y et laisser libres les déplacements suivant Z et les rotations suivant X, Y et Z.

Encastrement	
Raideur de rotation suivant X (N.m^{-1})	10^5
Raideur de rotation suivant Y (N.m^{-1})	$3,3.10^{-5}$
Bielle	
Longueur (mm)	460
Section (m^2)	$1,26.10^{-3}$
Module de Young (GPa)	200
Inertie de flexion (m^4)	$1,26.10^{-10}$
Inertie de torsion (m^4)	$2,52.10^{-10}$
Guide	
Raideur de rotation suivant Y (N.m^{-1})	$1,1.10^5$

TABLE 4.2 – Caractéristiques des liaisons.

De plus :

- la tuyauterie, d'une masse totale de l'ordre de 383kg, est pourvue d'une vanne de 120kg dont le centre de gravité est situé à 142,2mm de l'axe de la tuyauterie ;
- pour les calculs, la pression, égale au maximum à 95Bar, a été négligée après avoir vérifié que son influence était négligeable sur les comportements globaux linéaire et non-linéaire. La ligne était toutefois remplie d'eau lors des essais ;

- les essais dynamiques ont montré que le taux d’amortissement du 1^{er} mode de la maquette était de 1%.

Comportement et modes propres

Les trois premiers modes propres de la ligne sont les suivants :

- le 1^{er} mode propre se situe à 5Hz : il correspond à un mouvement global de la lyre dans son plan. Ce mode est piloté par le déplacement du point 35, et sollicite le coude situé près de l’encastrement en flexion hors plan ;
- le 2^e mode, prépondérant suivant Z, est à 5,9Hz. Celui-ci n’a pas été identifié lors des essais du rapport [7].
- le 3^e mode, piloté par le déplacement du point 19, est à 6,7Hz. C’est un mode d’ouverture-fermeture de la lyre dans son plan. Il sollicite les deux coudes de la lyre en ouverture-fermeture dans le plan de la lyre. Il correspond donc au second mode expérimental.

La figure 4.2 présente alors un exemple d’excitation, et la figure 4.3 compare les réponses linéaires à cette excitation pour le nœud 12 : selon un calcul complet (en rouge) et un calcul réalisé à partir des seules réponses des modes 1 et 3 (en bleu). Leur adéquation nous invite à ne considérer par la suite les réponses que suivant les trois premiers modes.

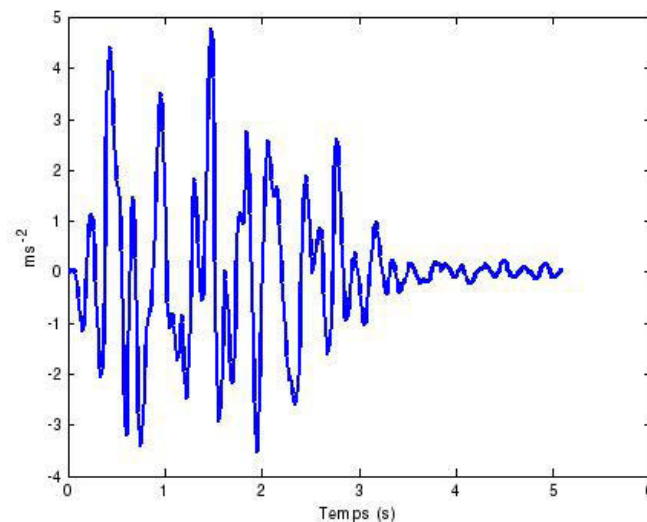


FIGURE 4.2 – *Exemple de signal d’excitation.*

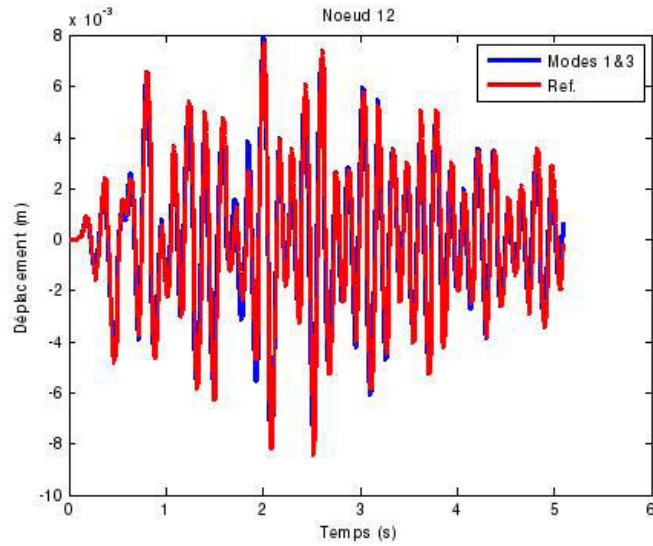


FIGURE 4.3 – *Comparaison, suivant le mode de calcul, de la réponse linéaire attachée au nœud 12.*

Les figures 4.4, 4.5 et 4.6 représentent alors les SRO à 1% des déplacements pour les trois différents nœuds indiqués : sont comparés le calcul complet (en rouge) et un calcul réalisé à partir des réponses des modes 1 et 3 uniquement (en bleu).

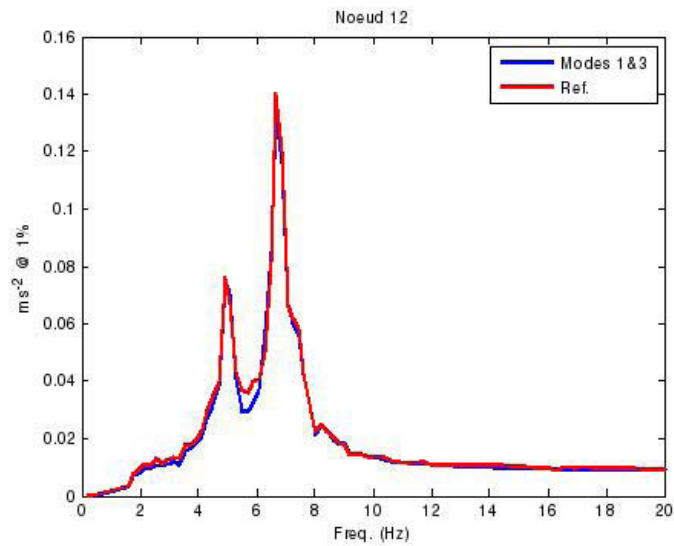


FIGURE 4.4 – *SRO à 1% du déplacement du nœud 12.*

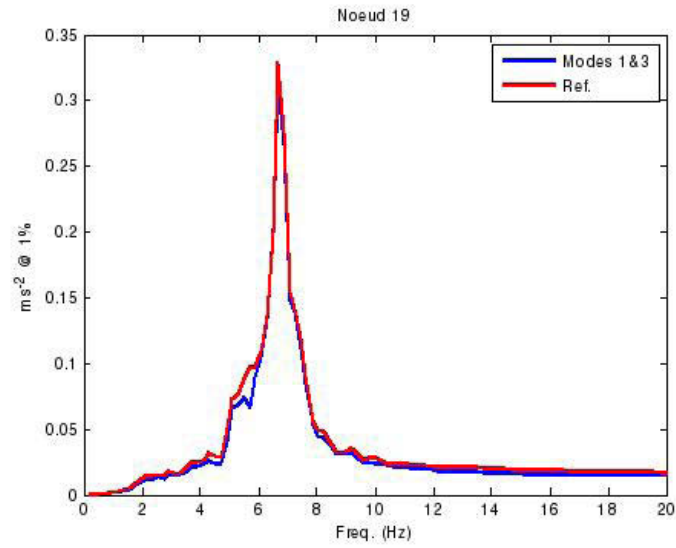


FIGURE 4.5 – SRO à 1% du déplacement du nœud 19.

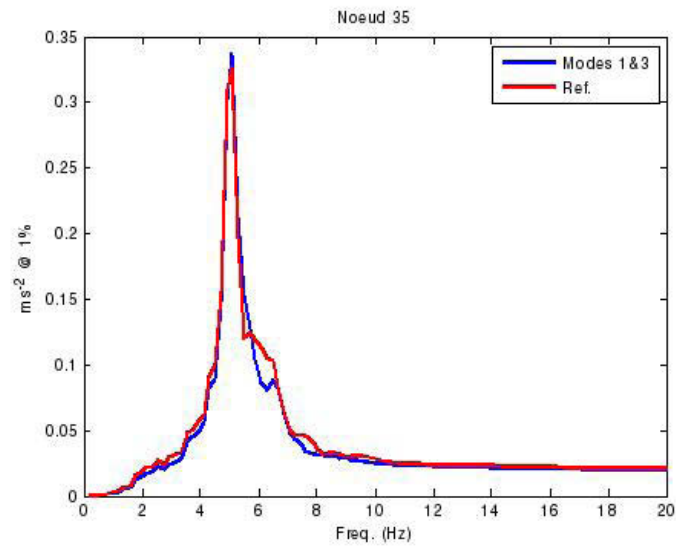


FIGURE 4.6 – SRO à 1% du déplacement du nœud 35.

Le déplacement du nœud 12 présente un caractère bimodal, puisqu'il fait intervenir les modes 1 et 3. Nous choisissons d'étudier ce nœud-ci, dans sa réponse linéaire, qui sera suffisante pour observer des éléments pathologiques.

4.2 Calcul de courbes de fragilité

Courbe-cible et ajustement log-normal

Les accélérogrammes utilisés sont alors, dans un premier temps, les 10 000 signaux artificiels de la famille E_1 , comme dans la partie validation du chapitre 3. Ils sont regroupés selon 20 classes de PGA. La structure choisie est soumise à tous ces signaux, fournissant donc 20 classes de signaux de réponse. Le critère de défaillance a été choisi empiriquement, de manière à obtenir des courbes suffisamment lisibles : $b = 2,5 \cdot 10^{-3} \text{m}$

Rappelons que nous examinons ici la réponse linéaire au niveau du nœud 12, en ne considérant que les trois premiers modes propres, conformément aux conclusions du paragraphe précédent.

La courbe de fragilité « cible » alors obtenue est présentée figure 4.7, avec son intervalle de confiance, ainsi que son peu convaincant ajustement log-normal.

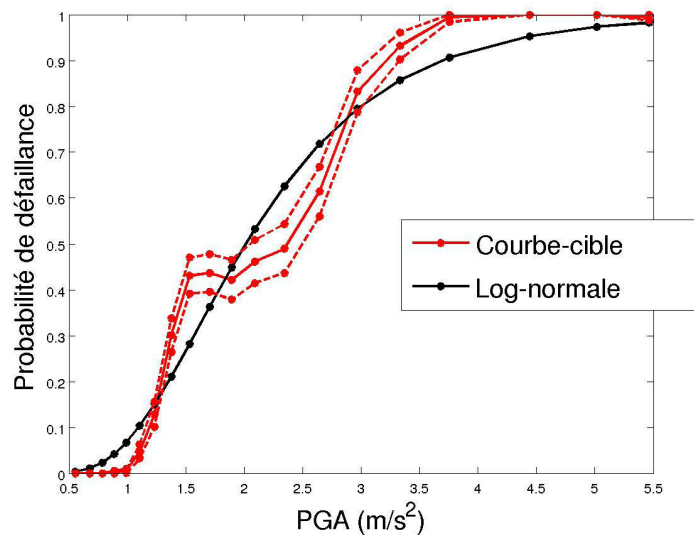


FIGURE 4.7 – Courbe de fragilité par classe de PGA pour le nœud 12 : réponses originelles et ajustement log-normal.

Enrichissement des classes de réponse

Procédons alors à l'enrichissement des classes de réponse. Dans chacune des classes de PGA, les signaux de réponse sont utilisés pour identifier un modèle K-L, avant de simuler 5 000 nouveaux signaux de réponse. A titre d'exemple, considérons la classe n°13 des signaux d'entrée, qui ne compte que 335 signaux. En premier lieu, la variance des deux familles successives de signaux est inchangée, comme en témoigne la figure 4.8.

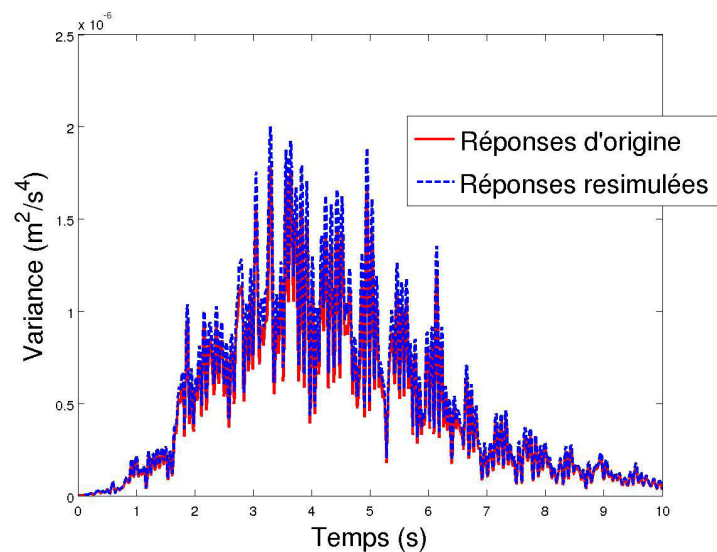


FIGURE 4.8 – Comparaison des variances pour les réponses d'origine et celles resimulées, pour la 13^e classe de réponse.

La figure 4.9 montre de même la concordance de la répartition des maxima.

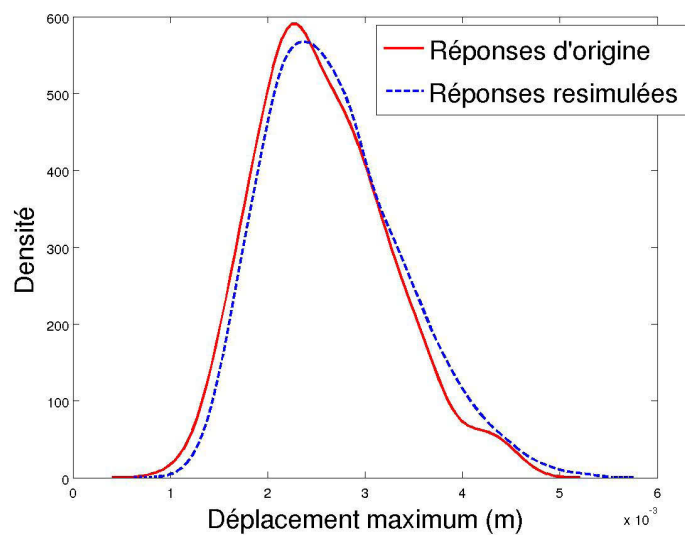


FIGURE 4.9 – Répartition des maxima des réponses d'origine et celles resimulées, pour la 13^e classe de réponse.

Et la figure 4.10 présente la bonne adéquation des deux répartition du processus à un instant fixé : celui du maximum de variance.

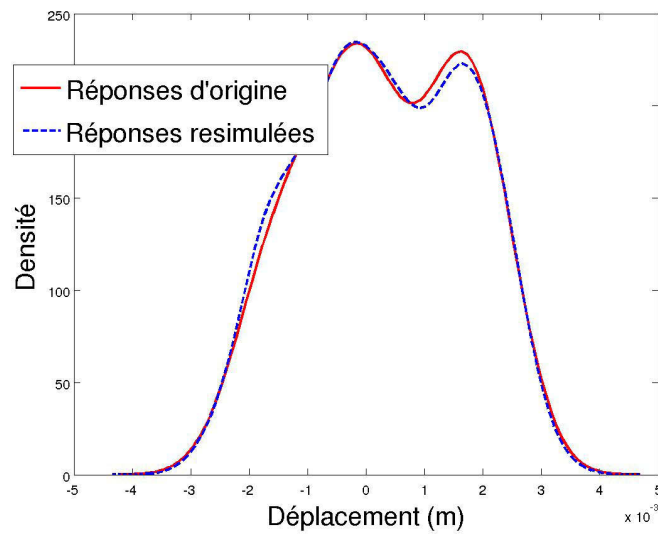


FIGURE 4.10 – Répartition, à l’instant de variance maximum, des réponses d’origine et celles resimulées, pour la 13^e classe de réponse.

Ces résultats invitent à poursuivre, et à tracer sur la figure 4.11 l’estimation de la courbe de fragilité obtenue avec ces signaux resimulés.

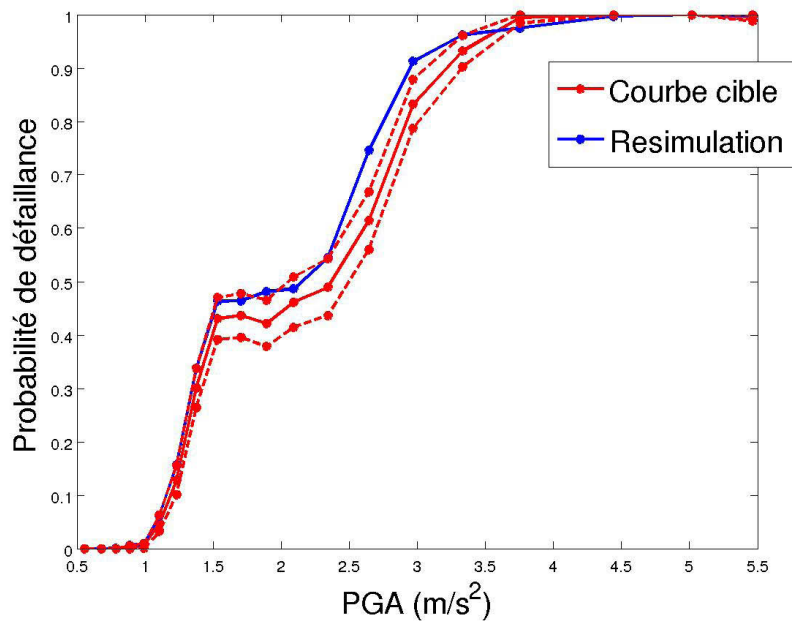


FIGURE 4.11 – Courbe de fragilité par classe de PGA, pour le nœud 12 : réponses originelles et resimulées.

Les conclusions sont proches de celles formulées à l’occasion de la validation du modèle. Le résultat

obtenu est satisfaisant puisque, une fois encore, la nouvelle courbe se place au-dessus de la courbe-cible, sauf pour les classes de forte intensité sismique.

Incertitudes : resimulation de la matrice R

Aus paragraphes 2.5.3 et 3.5, nous proposons de construire un intervalle autour de la courbe de fragilité pour prendre en compte l'incertitude de modèle, grâce à une resimulation de la matrice R . La figure 4.12 compare en premier lieu :

- en rouge et bleu, les courbes construites précédemment : la cible et la courbe issue des signaux resimulés ;
- en noir, l'intervalle construit lorsque l'intégralité des réponses de chaque classe est utilisée pour resimuler les matrices R .

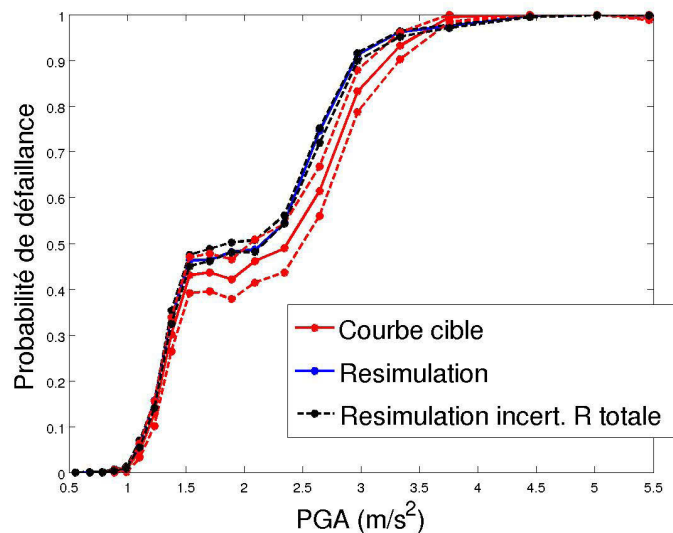


FIGURE 4.12 – Prise en compte de l'incertitude sur la matrice R : utilisation de la **totalité** des signaux de chaque classe de réponse.

La bonne adéquation des résultats confirme les conclusions du § 3.5.2 : la méthode consistant à « resimuler » la matrice R est valide, lorsque tous les signaux de réponse sont considérés. La figure 4.13 s'autorise alors à comparer :

- en rouge et noir, les courbes précédentes de mêmes couleurs ;
- en magenta, l'intervalle construit lorsque seules 30 réponses de chaque classe sont utilisées pour resimuler les matrices R .

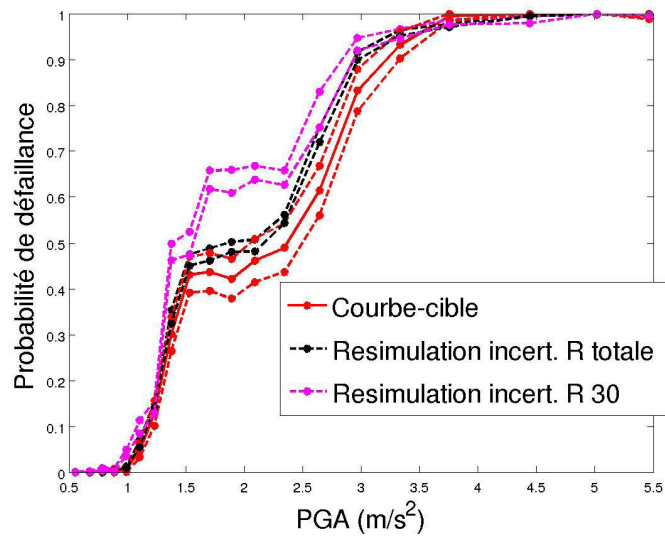


FIGURE 4.13 – *Prise en compte de l'incertitude sur la matrice R : utilisation de seulement 30 signaux par classe de réponse.*

L'intervalle de confiance magenta obtenu est très différent de la courbe-cible, situation déjà observée au chapitre 3. Mais notre discussion sur la courbe d'aléa de la page 70 nous invite à ne pas considérer les courbes de fragilité au-delà d'une certaine abscisse d'intensité sismique. Pour les classes 1 à 9, la concordance est cette fois très bonne comme le montre l'agrandissement de la figure 4.14, tout en conservant un léger caractère conservatif.

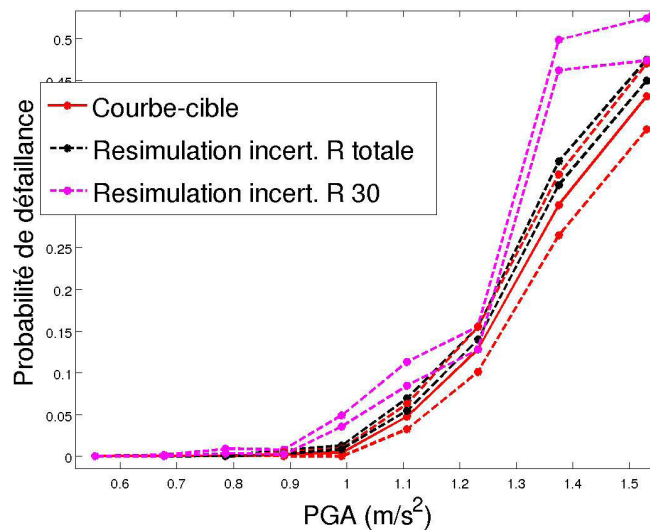


FIGURE 4.14 – *Agrandissement de la figure 4.13.*

4.3 Conclusion

Notre méthode, validée au paragraphe 3 sur des modèles de structures simples, nous semble également convaincante pour une structure issue d'un contexte industriel, à condition de ne s'intéresser qu'aux classes de faible intensité sismique. Pour ces classes, les résultats issus de la resimulations de la matrice R ne sont pas en parfaite concordance avec la courbe-cible, mais ils présentent néanmoins un caractère conservatif, susceptible d'intéresser l'ingénieur.

Conclusion et perspectives

La courbe de fragilité d'une structure sous séisme représente sa probabilité de défaillance en fonction de l'intensité sismique à laquelle elle est soumise. Ce mémoire s'est proposé d'y parvenir au moyen du modèle proposé par l'article [53]. Ce procédé se fonde sur la représentation de Karhunen-Loève pour simuler une grande quantité d'accélérogrammes artificiels à partir d'une famille d'accélérogrammes authentiques, choisis pour leur adéquation à un scénario sismique donné. Il a été possible de constater que prendre en compte la dépendance mutuelle des coefficients du développement améliore le résultat. Ces signaux une fois générés, puis triés par classes d'indicateur de nocivité sismique, un modèle à éléments finis de la structure permet d'obtenir des classes de signaux de réponses, puis une estimation de la courbe de fragilité de la structure. Cette méthode présente plusieurs atouts qui ont été soulignés au chapitre 3 au moment de la vérification de sa validité. La méthode est en premier lieu opérationnelle quel que soit l'indicateur de nocivité sismique utilisé. Les différentes caractéristiques du scénario sismique sont de plus conservées lors de la « resimulation », ce qui permet d'entamer des calculs de Monte-Carlo, même lorsque peu de signaux étaient disponibles à l'origine. Le processus stochastique modélisant le mouvement du sol est de plus supposé de la forme la plus générale possible, sans aucune hypothèse de gaussianité ou de stationnarité. Il est enfin possible de se dispenser de deux procédés courants : l'hypothèse que les courbes de fragilité suivent une loi log-normale, et la méthode du *scaling*, qui supposerait que le contenu spectral de la réponse d'une structure linéaire est proportionnelle à l'intensité de la sollicitation sismique.

Mais l'intérêt de ce procédé réside également dans deux améliorations de notre cru. Il est possible, et même souhaitable, d'enrichir également la famille des signaux de réponse. Ceci permet d'effectuer moins de calculs de structures, tout en disposant *in fine* de nombreux signaux de réponses, permettant donc une bonne évaluation de la courbe de fragilité. Mais ceci augmente d'autant plus l'incertitude sur une certaine matrice R , qui intervient de façon capitale dans la représentation de Karhunen-Loève. Nous proposons une prise en compte « partielle » de cette incertitude lorsque les signaux de réponses disponibles sont peu nombreux, permettant alors une estimation convenable de la courbe même lorsque leur nombre est très réduit.

Nous concluons en listant les différents problèmes qui furent soulevés au cours de ce travail, et qui pourraient faire l'objet d'études ultérieures. Nous avons exploré, au chapitre 2, la possibilité de s'appuyer sur une base de polynômes orthogonaux pour mieux exprimer des densités de probabilité, et cette exploration mérite d'être poursuivie. Notre méthode présente de plus l'inconvénient de simuler beaucoup de signaux identiques, ce qui limite le nombre de simulations possibles ; à moins de ne considérer la courbe qu'au

niveau des faibles intensités sismiques, pour lesquelles le problème n'apparaît pas. Mais le point principal reste le fait que notre méthode n'aboutisse, presque tout le temps, qu'à des courbes situées *au dessus* des courbes cibles ; ceci, que le procédé soit utilisé seul, combiné avec une resimulation des signaux de réponses ou même avec la prise en compte de l'incertitude sur la fameuse matrice R .

Annexe A

Rappels de probabilités réelles

Rappelons ici les notions et résultats essentiels à la lecture de ce mémoire. Pour des ouvrages détaillés, nous conseillons [21] ou [31].

Soit (Ω, \mathcal{T}, P) un espace probabilisé.

Si \mathcal{B} est la tribu borélienne de \mathbb{R} , dt la mesure de Lebesgue sur \mathbb{R} , alors $(\mathbb{R}, \mathcal{B}, dt)$ est un espace mesuré.

Pour tout $N \geq 2$, si \mathcal{B}^N est la tribu borélienne de \mathbb{R}^N , $d\lambda_N$ la mesure de Lebesgue de \mathbb{R}^N , abusivement notée $dt_1 \dots dt_N$ ou simplement $d\mathbf{t}$, alors $(\mathbb{R}^N, \mathcal{B}^N, d\mathbf{t})$ est un espace mesuré.

A.1 Variables aléatoires

Une variable aléatoire (v.a.) réelle X est une fonction mesurable de (Ω, \mathcal{T}) vers $(\mathbb{R}, \mathcal{B})$.

Les notions de fonction de répartition, d'espérance, de variance, de densité et de fonction caractéristique sont supposées connues du lecteur. Bornons-nous à rappeler pour mémoire l'expression de la densité de probabilité :

- de la **loi normale** $\mathcal{N}(\mu, \sigma^2)$ avec $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$. Elle vérifie $\forall x \in \mathbb{R}$:

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2}$$

- de la **loi log-normale** $\mathcal{LN}(A_m, \beta)$ avec $(A_m, \beta) \in (\mathbb{R}_+^*)^2$. Elle vérifie $\forall x \in \mathbb{R}$:

$$p(x) = \frac{1}{x\beta \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(x/A_m)}{\beta} \right)^2}$$

Rappelons de plus que :

- une v.a. X est dite d'ordre $r > 0$ si $E[|X|^r] < +\infty$;
- deux v.a. X et Y sont dites décorrélées ou orthogonales si $E[XY] = 0$.

Convergence d'une suite de v.a.

Une suite de v.a. réelles $(X_n)_{n \geq 1}$ **converge en loi** vers une v.a. X si la suite de ses fonctions de répartition converge simplement vers la fonction de répartition de X en tous ses points de continuité ; ce qui se note :

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$$

Une suite de v.a. réelles $(X_n)_{n \geq 1}$ **converge en moyenne d'ordre $r > 0$** vers une v.a. X si :

$$\lim_{n \rightarrow +\infty} E[|X_n - X|^r] = 0$$

ce qui se note :

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{L}^r} X$$

Dans le cas $r = 2$, on parle de **convergence en moyenne quadratique**.

Théorème central limite

Soit $(X_n)_{n \geq 1}$ une suite de v.a. réelles, indépendantes et identiquement distribuées (i.i.d.), de même espérance μ et de même variance σ^2 .

Si $\forall n \geq 1, \bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$, alors :

$$\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

A.2 Vecteurs aléatoires

Un vecteur aléatoire réel $\mathbf{X} = (X_1, \dots, X_N)$ est une application mesurable de (Ω, \mathcal{T}) vers $(\mathbb{R}^N, \mathcal{B}^N)$. Sa fonction de répartition $F_{\mathbf{X}}$ se définit comme :

$$F_{\mathbf{X}} : \begin{cases} \mathbb{R}^N & \rightarrow \mathbb{R} \\ \mathbf{x} = (x_1, \dots, x_N) & \rightarrow P\left(\bigcap_{i=1}^N (X_i \leq x_i)\right) \end{cases}$$

Si $p_{\mathbf{X}}$ est sa densité de probabilité conjointe vérifiant, $\forall \mathbf{x} = (x_1, \dots, x_N)$:

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_N} p_{\mathbf{X}}(\mathbf{t}) d\mathbf{t}$$

alors $\forall i \in \llbracket 1; N \rrbracket$ la probabilité marginale p_{X_i} de la v.a. réelle X_i s'écrit

$$p_{X_i} : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ u & \rightarrow \int_{\mathbb{R}^{M-1}} p_X(t_1, \dots, t_{i-1}, u, t_{i+1}, \dots, t_N) dt_1 \dots dt_{i-1} dt_{i+1} \dots dt_N \end{cases}$$

Comme précédemment, rappelons la densité de probabilité d'une loi multinormale $\mathcal{N}(\mu, \Gamma)$, paramétrée par un « vecteur moyenne » $\mu \in \mathbb{R}^N$ et une « matrice de variance-covariance » $\Gamma \in \mathcal{S}_N^{+\star}(\mathbb{R})$. Un vecteur X qui suit cette distribution admet la densité de probabilité suivante :

$$p_X : \begin{cases} \mathbb{R}^N & \rightarrow \mathbb{R} \\ \mathbf{x} = (x_1, \dots, x_N) & \rightarrow \frac{1}{\sqrt{(2\pi)^N |\Gamma|}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^\top \Gamma^{-1}(\mathbf{x}-\mu)} \end{cases}$$

Sa simulation s'effectue comme suit. Puisque $\Gamma \in \mathcal{S}_N^{+\star}(\mathbb{R})$, la factorisation de Cholesky assure l'existence de $C \in \mathcal{M}_N(\mathbb{R})$ telle que $\Gamma = CC^\top$. En notant $U \hookrightarrow \mathcal{N}(\mathbf{0}, I_N)$ qui est aisément simulable (il s'agit d'un vecteur à N composantes gaussiennes centrées réduites indépendantes), alors :

$$X \stackrel{\mathcal{L}}{=} CU + \mu \quad (\text{égalité en loi})$$

Théorème central limite

Soit $(X_n)_{n \geq 1}$ une suite de vecteurs aléatoires réels, indépendants et identiquement distribués (i.i.d.), de même moyenne μ et de même matrice de variance-covariance Γ .

Si $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$, alors :

$$\boxed{\frac{\bar{X}_n - \mu}{1/\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}, \Gamma)}$$

A.3 Processus stochastiques

A.3.1 Généralités

Soit A un sous-ensemble de \mathbb{R} . Dans ce mémoire, l'ensemble A désigne soit \mathbb{R} tout entier, soit un intervalle $[0, T]$ avec $T > 0$.

Un processus stochastique réel $X = \{X_t\}_{t \in A}$ est une famille de v.a. réelles indexée par A .

Par la suite, on notera abusivement $X = \{X(t)\}_{t \in A}$.

On supposera n'avoir affaire qu'à des processus stochastiques à densité, c'est-à-dire qu'il existe une fonc-

tion $p_X \in \mathcal{C}^0(\mathbb{R} \times A, [0,1])$ vérifiant, $\forall t \in \mathbb{R}$:

$$P(X(t) \leq x) = \int_{\mathbb{R}} p_X(x, t) dx$$

La **fonction de moyenne** du processus μ_X du processus $X = \{X(t)\}_{t \in A}$ est définie comme :

$$\mu_X : \begin{cases} A & \rightarrow \mathbb{R} \\ t & \rightarrow E[X(t)] = \int_{\mathbb{R}} x p_X(x, t) dx \end{cases}$$

Sa **fonction de variance** V_X est définie comme :

$$V_X : \begin{cases} A & \rightarrow \mathbb{R} \\ t & \rightarrow E[(X(t) - \mu_X)^2] = E[X(t)^2] - \mu_X(t)^2 \end{cases}$$

Et sa **fonction d'autocorrélation** R_X comme :

$$R_X : \begin{cases} A \times A & \rightarrow \mathbb{R} \\ (t, t') & \rightarrow E[X(t)X(t')] \end{cases}$$

Le processus $X = \{X(t)\}_{t \in A}$ est dit **centré** si $\forall t \in A, \mu_X(t) = 0$.

Lorsque c'est le cas, $\forall (t, t') \in A : V_X(t) = R_X(t, t)$

Par ailleurs, pour $r > 0$, un processus $X = \{X(t)\}_{t \in \mathbb{R}}$ est dit :

— d'ordre r si $\forall t \in A$

$$E[|X(t)|^r] < +\infty$$

— continu en moyenne d'ordre r si $\forall t \in A$

$$\lim_{t' \rightarrow t} E[|X(t) - X(t')|^r] = 0$$

En particulier, « continu en moyenne d'ordre deux » est couramment noté « continu en m.o.d. ».

A.3.2 Processus stationnaires

Un processus $X = \{X(t)\}_{t \in \mathbb{R}}$ indexé sur \mathbb{R} est dit **strictement stationnaire** si $\forall u \in \mathbb{R}, \{X(t+u)\}_{t \in \mathbb{R}}$ a même loi que $\{X(t)\}_{t \in \mathbb{R}}$. Dans ce cas, la fonction d'autocorrélation ne dépend que de la différence $t - t'$ et **se renote abusivement** R_X :

$$R_X(u) = E[X(0)X(0+u)]$$

avec cette fois-ci R_X une fonction de \mathbb{R} dans \mathbb{R} .

Alors si X est centré, stationnaire et continu en m.o.d., il existe dM_X ([73], p. 343) une mesure positive bornée sur \mathbb{R} vérifiant :

$$R_X(\tau) = \int_{\mathbb{R}} e^{i\omega\tau} dM_X(\omega)$$

Si la mesure spectrale dM_X admet une densité S_X par rapport à la mesure de Lebesgue $d\omega$:

$$dM_X(\omega) = S_X(\omega) d\omega$$

La fonction S_X est alors appelée **Densité Spectrale de Puissance (DSP, [56])** du processus X . La fonction d'autocorrélation de X s'écrit :

$$R_X(\tau) = \int_{-\infty}^{+\infty} S_X(\omega) e^{i\omega\tau} d\omega$$

Le théorème de Wiener-Khintchine (voir [73]) précise enfin que, si $R_X \in L^1_{\mathbb{R}}(\mathbb{R}, dt)$, alors :

$$S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_X(\tau) e^{-i\omega\tau} d\tau$$

Des outils plus sophistiqués ¹ permettraient alors de définir rigoureusement le **bruit blanc** : un processus de DSP constante et de fonction d'autocorrélation correspondant à une impulsion de Dirac.

1. Comme la théorie des distributions.

Annexe B

Echantillonnage

La réalisation de calculs numériques impose un échantillonnage des signaux, dans le domaine temporel et le domaine fréquentiel.

Pour échantillonner un signal de durée T en un vecteur à N éléments, le pas de temps sera $\Delta t = \frac{T}{N-1}$. Le « vecteur temps » discret est :

$$\{t_k = (k-1)\Delta t\}_{k \in \llbracket 1;N \rrbracket}$$

Dans le domaine fréquentiel, la pulsation est bornée par une « pulsation de coupure » Ω_c qu'il convient de fixer : ainsi $\omega \in [-\Omega_c, \Omega_c]$, intervalle de longueur $2\Omega_c$, d'où un pas de pulsation :

$$\Delta\omega = \frac{2\Omega_c}{N}$$

Si $\Omega_e = \frac{2\pi}{\Delta t}$ est la « pulsation d'échantillonnage » en temps, le théorème de Shannon [5] majore la pulsation de coupure autorisée : $\Omega_c \leq \frac{\Omega_e}{2}$. Nous choisissons $\Omega_c = \frac{\Omega_e}{2}$.
A supposer que $N \in 2\mathbb{N}$, le « vecteur pulsation » discret de longueur N est :

$$\left\{ \omega_p = (2p+1)\frac{\Delta\omega}{2} \right\}_{p \in \llbracket -\frac{N}{2}; \frac{N}{2}-1 \rrbracket} = \left\{ \omega_p = -\Omega_c + (2p-1)\frac{\Delta\omega}{2} \right\}_{p \in \llbracket 1;N \rrbracket}$$

Dans ce mémoire, les calculs sont effectués en pulsation, mais les figures tracées en fréquence, comme c'est l'usage dans les applications industrielles.

Condition sur Δt

Nous considérerons par la suite un oscillateur linéaire à un degré de liberté de paramètres (ω_0, ξ_0) soumis à un accélérogramme $\{\gamma(t)\}_{t \in [0, T]}$. Sa réponse en déplacement y vérifie l'équation :

$$\ddot{y}(t) + 2\xi_0\omega_0\dot{y}(t) + \omega_0^2 y(t) = -\gamma(t)$$

Si on pose :

$$Y(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix} \quad A = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\omega_0\xi_0 \end{pmatrix} \quad F(t) = \begin{pmatrix} 0 \\ -\gamma(t) \end{pmatrix}$$

alors il vient que :

$$\dot{Y}(t) = AY(t) + F(t) \stackrel{\text{noté}}{=} f(t, Y(t))$$

On cherchera à déterminer des valeurs approchées Y_1, \dots, Y_N des valeurs $Y(t_k)$ prises par la solution exacte Y , grâce à la méthode numérique « d'Euler explicite » de pas Δt [14].

Le schéma numérique est le suivant. $\forall k \in \llbracket 1; N \rrbracket$:

$$\begin{aligned} Y_{k+1} &= Y_k + \Delta t f(t_k, Y_k) \\ &= Y_k + \Delta t (AY_k + F(t_k)) \\ &= (I + \Delta t A) Y_k + \Delta t F(t_k) \end{aligned}$$

et le rayon spectral de la matrice $I + \Delta t A$ doit être strictement plus petit que 1 :

$$\rho(I + \Delta t A) = \max(|\lambda|, \lambda \in \mathcal{S}_{\mathbb{R}}(I + \Delta t A)) < 1$$

Comme :

$$I + \Delta t A = \begin{pmatrix} 1 & \Delta t \\ -\omega_0^2 \Delta t & 1 - 2\omega_0\xi_0 \Delta t \end{pmatrix}$$

son polynôme caractéristique s'écrit :

$$\chi_{I + \Delta t A}(X) = X^2 - 2(1 - \omega_0\xi_0 \Delta t)X + (1 - 2\omega_0\xi_0 \Delta t + \omega_0^2 \Delta t^2)$$

de racines conjuguées :

- $x_1 = 1 - \omega_0\xi_0 \Delta t + i\omega_0 \sqrt{1 - \xi_0^2 \Delta t}$
- $x_2 = 1 - \omega_0\xi_0 \Delta t - i\omega_0 \sqrt{1 - \xi_0^2 \Delta t}$

ainsi :

$$\begin{aligned} \rho(I + \Delta t A) &= |x_1| = |x_2| \\ &= \sqrt{(1 - \omega_0\xi_0 \Delta t)^2 + \omega_0^2(1 - \xi_0^2 \Delta t)^2} \\ &= \sqrt{1 - 2\omega_0\xi_0 \Delta t + \omega_0^2 \Delta t^2} \end{aligned}$$

donc :

$$\begin{aligned}
& \rho(I + \Delta t A) < 1 \\
& \Leftrightarrow \sqrt{1 - 2\omega_0 \xi_0 \Delta t + \omega_0^2 \Delta t^2} < 1 \\
& \Leftrightarrow \omega_0^2 \Delta t (\Delta t - 2 \frac{\xi_0}{\omega_0}) < 0 \\
& \Leftrightarrow \boxed{\Delta t \leq 2 \frac{\xi_0}{\omega_0}}
\end{aligned}$$

Au chapitre 3, nous manipulons un oscillateur linéaire de paramètres $\omega_0 = 2\pi.5 \text{ rad.s}^{-1}$ et $\xi_0 = 0.05$, ce qui exige $\Delta t \leq 2 \frac{\xi_0}{\omega_0} = 3,2.10^{-3} \text{ s}$. Les signaux dont nous disposons sont enregistrés avec un pas de temps de 10^{-2} s , ce qui est insuffisant et impose leur ré-échantillonnage.

Ré-échantillonnage

Le ré-échantillonnage d'un signal $\{x(t)\}_{t \in [0, T]}$ de pas de temps Δt en un signal de pas de temps plus petit $\Delta t'$ s'effectue alors de la façon suivante :

- $N' = \frac{T}{\Delta t'} + 1$ est le nombre de points du signal ré-échantillonné ;
- $\Omega'_c = \frac{N'}{N} \Omega_c$ devient la nouvelle pulsation de coupure et $\left\{ \omega'_p = (2p + 1) \frac{\Delta \omega}{2} \right\}_{p \in \llbracket -\frac{N'}{2}; \frac{N'}{2} - 1 \rrbracket}$ le nouveau « vecteur pulsation » discret ;
- soit $F : \omega'_p \rightarrow \hat{x}(\omega_p)$ si $-\frac{N}{2} \leq p \leq \frac{N}{2} - 1$ et 0 sinon ;
- le signal ré-échantillonné est la transformée de Fourier discrète inverse \check{F} , de longueur N' .

Pour les valeurs que nous avons choisies, un ré-échantillonnage vers un nouveau de pas de temps $\Delta t' = 10^{-3} \text{ s}$ est suffisant.

Annexe C

Démonstration du théorème de Karhunen-Loève

H est toujours l'espace de Hilbert $L^2_{\mathbb{R}}([0, T], dt)$ muni de son produit scalaire usuel :

$$(f|g) = \int_0^T f(t)g(t) dt \quad (\text{C.1})$$

Enoncé

Soit $X = \{X(t)\}_{t \in [0, T]}$ un processus stochastique réel défini sur un espace probabilisé (Ω, \mathcal{T}, P) et indexé sur $[0, T]$, du second ordre, centré, continu en m.o.d.¹, à trajectoires continues. Alors il existe :

- $\{\phi_\alpha\}_{\alpha \geq 1}$ une base hilbertienne de $(H, (\cdot|\cdot))$;
- $\{\lambda_\alpha\}_{\alpha \geq 1}$ une famille de réels positifs, ordonnés de façon décroissante ;
- $\{\xi_\alpha\}_{\alpha \geq 1}$ une famille dénombrable de v.a. réelles définies sur (Ω, \mathcal{T}, P) , **centrées, de variance 1, et deux à deux orthogonales (i.e. non corrélées)** ;

vérifiant, $\forall t \in [0, T]$:

$$X(t) = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t) \quad (\text{C.2})$$

Cette représentation s'entend au sens **uniforme en moyenne quadratique**, ce qui signifie que :

$$\lim_{N \rightarrow +\infty} \sup_{t \in [0, T]} E \left[\left\| X(t) - \sum_{\alpha=1}^N \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t) \right\|^2 \right] = 0 \quad (\text{C.3})$$

1. Voir p.90

Les v.a. $\{\xi_\alpha\}_{\alpha \geq 1}$ s'obtiennent de plus selon :

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \int_0^T X(t) \phi_\alpha(t) dt \quad (\text{C.4})$$

Démonstration

Soit $R_X : (t, t') \rightarrow E[X(t)X(t')]$ la fonction d'autocorrélation du processus X . L'opérateur Q_X est défini comme :

$$Q_X : \begin{cases} H & \rightarrow H \\ f & \rightarrow Q_X f : \begin{cases} [0, T] & \rightarrow \mathbb{R} \\ t & \rightarrow \int_0^T R_X(t, t') f(t') dt' \end{cases} \end{cases}$$

La fonction R_X est continue, symétrique et de type positif ce qui fait d'elle un « noyau de Mercer »². Alors d'après le **théorème de Mercer**² :

1. le spectre (dénombrable) de Q_X est inclus dans \mathbb{R}_+ . Il est noté $\{\lambda_\alpha\}_{\alpha \geq 1}$, ordonné de façon décroissante ;
2. il existe $\{\phi_\alpha\}_{\alpha \geq 1}$ une base hilbertienne de H constituée de fonctions propres de Q_X associées aux valeurs propres $\{\lambda_\alpha\}_{\alpha \geq 1}$;
3. les fonctions propres correspondant à des valeurs propres strictement positives sont continues sur $[0, T]$;
4. R_X admet la représentation (somme uniformément convergente sur $[0, T]^2$) :

$$R_X(t, t') = \sum_{\alpha=1}^{+\infty} \lambda_\alpha \phi_\alpha(t) \phi_\alpha(t') \quad (\text{C.5})$$

Par ailleurs, puisque $\{\phi_\alpha\}_{\alpha \geq 1}$ est une base hilbertienne de H :

$$X = \sum_{\alpha=1}^{+\infty} (X|\phi_\alpha) \phi_\alpha \quad (\text{C.6})$$

Alors $\forall \alpha \geq 1$, on définit :

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} (X|\phi_\alpha) = \frac{1}{\sqrt{\lambda_\alpha}} \int_0^T X(t) \phi_\alpha(t) dt \quad (\text{C.7})$$

ce qui amène à :

$$X = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha \quad (\text{C.8})$$

2. La démonstration [47], difficile, de ce résultat repose sur la compacité de l'opérateur Q_X , obtenue grâce au théorème d'Ascoli.

Propriétés des ξ_α

En premier lieu, $\forall \alpha \geq 1$:

$$\begin{aligned} E[\xi_\alpha] &= E\left[\frac{1}{\sqrt{\lambda_\alpha}} \int_0^T X(t)\phi_\alpha(t) dt\right] \\ &= \frac{1}{\sqrt{\lambda_\alpha}} \int_0^T E[X(t)] \phi_\alpha(t) dt \\ &= 0 \quad \text{car } X \text{ est centré.} \end{aligned}$$

Donc ξ_α est **centrée**.

Ensuite, $\forall \alpha$ et $\beta \geq 1$:

$$\begin{aligned} E[\xi_\alpha \xi_\beta] &= E\left[\frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \int_0^T X(t)X(t')\phi_\alpha(t)\phi_\beta(t') dt dt'\right] \\ E[\xi_\alpha \xi_\beta] &= \frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \int_0^T E[X(t)X(t')] \phi_\alpha(t)\phi_\beta(t') dt dt' \\ &= \frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \int_0^T R_X(t, t') \phi_\alpha(t)\phi_\beta(t') dt dt' \\ &= \frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \phi_\beta(t') \left(\int_0^T R_X(t, t') \phi_\alpha(t) dt \right) dt' \\ &= \frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \phi_\beta(t') \times Q_X(\phi_\alpha)(t') dt' \\ &= \frac{1}{\sqrt{\lambda_\alpha \lambda_\beta}} \int_0^T \phi_\beta(t') \times \lambda_\alpha \phi_\alpha(t') dt' \quad \text{car } \phi_\alpha \text{ est vecteur propre de } Q_X \\ &= \sqrt{\frac{\lambda_\alpha}{\lambda_\beta}} \int_0^T \phi_\beta(t') \phi_\alpha(t') dt' \\ &= \sqrt{\frac{\lambda_\alpha}{\lambda_\beta}} (\phi_\beta | \phi_\alpha) \\ &= \sqrt{\frac{\lambda_\alpha}{\lambda_\beta}} \delta_{\alpha\beta} \quad \text{car } \{\phi_\alpha\}_{\alpha \geq 1} \text{ est orthonormée} \end{aligned} \tag{C.9}$$

ce qui implique bien que les $\{\xi_\alpha\}_{\alpha \geq 1}$ sont **décorrélées** et de **variance 1**.

Convergence de la somme

L'égalité fonctionnelle $X = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha$ n'est valable que d'un point de vue hilbertien, et son élargissement pour tout t (en particulier la convergence uniforme de la somme) nécessite d'être démontré. On note

donc $\forall(N,t) \in \mathbb{N}^\star \times [0,T]$:

$$S_N(t) = \sum_{\alpha=1}^N \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t) = \sum_{\alpha=1}^N \left(\int_0^T X(s) \phi_\alpha(s) ds \right) \phi_\alpha(t) \quad (\text{C.10})$$

Il vient alors que $\forall(N,t) \in \mathbb{N}^\star \times [0,T]$:

$$\begin{aligned} E[|X(t) - S_N(t)|^2] &= E[X(t)^2] + E[S_N(t)^2] - 2E[X(t)S_N(t)] \\ &= R_X(t,t) + E[S_N(t)^2] - 2E[X(t)S_N(t)] \end{aligned}$$

D'une part :

$$\begin{aligned} E[S_N(t)^2] &= E \left[\sum_{\alpha=1}^N \sum_{\beta=1}^N \sqrt{\lambda_\alpha} \sqrt{\lambda_\beta} \xi_\alpha \xi_\beta \phi_\alpha(t) \phi_\beta(t) \right] \\ &= \sum_{\alpha=1}^N \sum_{\beta=1}^N \sqrt{\lambda_\alpha \lambda_\beta} E[\xi_\alpha \xi_\beta] \phi_\alpha(t) \phi_\beta(t) \\ &= \sum_{\alpha=1}^N \sum_{\beta=1}^N \sqrt{\lambda_\alpha \lambda_\beta} \sqrt{\frac{\lambda_\alpha}{\lambda_\beta}} \delta_{\alpha\beta} \phi_\alpha(t) \phi_\beta(t) \\ &= \sum_{\alpha=1}^N \sum_{\beta=1}^N \lambda_\alpha \delta_{\alpha\beta} \phi_\alpha(t) \phi_\beta(t) \\ &= \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 \end{aligned} \quad (\text{C.11})$$

D'autre part :

$$\begin{aligned} E[X(t)S_N(t)] &= E \left[X(t) \sum_{\alpha=1}^N \left(\int_0^T X(s) \phi_\alpha(s) ds \right) \phi_\alpha(t) \right] \\ &= \sum_{\alpha=1}^N \left(\int_0^T E[X(t)X(s)] \phi_\alpha(s) ds \right) \phi_\alpha(t) \\ &= \sum_{\alpha=1}^N \left(\int_0^T R_X(t,s) \phi_\alpha(s) ds \right) \phi_\alpha(t) \\ &= \sum_{\alpha=1}^N Q_X(\phi_\alpha)(t) \phi_\alpha(t) \\ &= \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t) \times \phi_\alpha(t) \\ &= \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 \end{aligned}$$

Ainsi, $\forall(N,t) \in \mathbb{N}^\star \times [0,T]$:

$$\begin{aligned} E \left[|X(t) - S_N(t)|^2 \right] &= R_X(t,t) + \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 - 2 \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 \\ &= R_X(t,t) - \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 \end{aligned} \quad (\text{C.12})$$

Or la convergence est uniforme dans la représentation (C.5), ce qui signifie que :

$$\lim_{N \rightarrow +\infty} \sup_{(t,t') \in [0,T]^2} \left| R_X(t,t') - \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t) \phi_\alpha(t') \right| = 0 \quad (\text{C.13})$$

et donc, *a fortiori*, que :

$$\lim_{N \rightarrow +\infty} \sup_{t \in [0,T]} \left| R_X(t,t) - \sum_{\alpha=1}^N \lambda_\alpha \phi_\alpha(t)^2 \right| = 0 \quad (\text{C.14})$$

ce qui implique bien la convergence uniforme en moyenne quadratique :

$$\lim_{N \rightarrow +\infty} \sup_{t \in [0,T]} E \left[|X(t) - S_N(t)|^2 \right] = 0$$

c'est-à-dire que $\forall t \in [0,T]$:

$$\boxed{X(t) = \sum_{\alpha=1}^{+\infty} \sqrt{\lambda_\alpha} \xi_\alpha \phi_\alpha(t)} \quad (\text{C.15})$$

Références

- [1] G Ahmadi and F.G. Fan. Nonstationary kanai-tajimi models for el centro 1940 and mexico city 1985 earthquake. *Probabilistic Engineering Mechanics*, 5 :171–181, 1990.
- [2] M Amin and AHS Ang. Nonstationary stochastic model of earthquake motions. *Journal of engineering mechanics division*, 94 :559–583, 1968. (ASCE).
- [3] J-M Azais and M Wschebor. *Level sets and extrema of random processes and fields*. Wiley, 2009.
- [4] M Barbato and J-P Conte. Spectral characteristics of non-stationary random processes : theory and applications to linear structural models. *Probabilistic Engineering mechanics*, 23 :416–426, 2008.
- [5] M Bergounioux. *Mathématiques pour le traitement du signal*. Dunod, 2010.
- [6] J. Bernardo and A.F.M. Smith. *Bayesian Theory*. John Wiley, New York, 1994.
- [7] N Blay. Synthèse des essais et calculs sismiques réalisés sur le tronçon de ligne asg. Rapport SEMT/EMSI/RT/98/002.
- [8] H Brezis. *Analyse fonctionnelle : théorie et applications*. Masson, 1983.
- [9] P Cacciola. A stochastic approach for generating spectrum compatible fully non stationary earthquakes. *Computers and structures*, 88, 2010.
- [10] P Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, 1988.
- [11] R Clough and J Penzien. *Dynamics of structures*. McGraw-Hill, 1975.
- [12] R-B Corotis, E-H Vanmarcke, and C-A Cornell. First passage of nonstationary random processes. *Journal of Engineering Mechanics Division ASME*, 98 :401–414, 1972.
- [13] S Da-Veiga. *Analyse d'incertitudes et de sensibilité. Application aux modèles de cinétique chimique*. Mathématiques, spécialité statistiques.
- [14] J-P Demailly. *Analyse numérique et équations différentielles*. EDP Sciences, 1990.
- [15] M Di Paola. Transient spectral moments of linear systems. *SM Archives*, 10 :225–243, 1985.
- [16] SL Dimova and P Negro. Seismic assessment of an industrial frame structure designed according to eurocodes. part 2 : capacity and vulnerability. *Eng. structures*, 27 :724–735, 2005.
- [17] B Ellingwood and K Kinali. Quantifying and communicating uncertainty in seismic risk assessment. *Structural safety*, 31(2) :179–189, 2009.

- [18] EPRI. *Applications guide for use of seismic margin assessment in quantitative risk-informed decision-making*, 2004. Final report 1010998.
- [19] ESD. European strong-motion database. http://www.isesd.hi.is/ESD_Local/frameset.htm.
- [20] C Feau, H Badi, P Bernard, F Bernardin, and M Fogli. *Bibliographie sur l'étude des maxima d'un processus aléatoire - application aux courbes de fragilité sismiques*, 2010. Projet ANR CISSI, contribution au sous-projet 1, rapport final.
- [21] D Foata and A Fuchs. *Processus stochastiques*. Dunod, 2004.
- [22] S Geman and D Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, 1984.
- [23] R Ghanem and S Dham. Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Porous Media*, 32 :239–262, 1998.
- [24] R Ghanem and Spanos P. *Stochastic finite elements : a spectral approach*. Springer-Verlag, 1991.
- [25] R-J Gibert. *Vibrations des structures - Interactions avec les fluides, Sources d'excitation aléatoires*. Collection de la Direction des études et recherches d'EDF, 1988.
- [26] F Giraudeau. Contribution à l'évaluation probabiliste de la fragilité sismique des centrales nucléaires - cr-lamsid-11-003, cr-lamsid-11-007, cr-lamsid-11-010, rapport de première année de thèse. Feb. 2012.
- [27] EM Güneyisi and G Altay. Seismic fragility assessment of effectiveness of viscous dampers in r/c buildings under scenario earthquake. *Structural safety*, 30(5) :461–480, 2008.
- [28] RO Hamburger, DA Foutch, and Cornell CA. Translating research to practice : Fema/sac performance-based design procedures. *Earthquake Spectra*, 19(2) :255–267, 2003.
- [29] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.
- [30] S J J Ilango, S Sarkar, and A Sameen. Reconstruction of 2-d porous media using karhunen-loève expansion. *Probabilistic Engineering Mechanics*, 32 :56–65, 2013.
- [31] J Jacod and P Protter. *L'essentiel en théorie des probabilités*. Cassini, 2003.
- [32] A.K. Jain, M.N. Murty, and P.J. Flynn. Data clustering : a review. *ACM Computing Surveys*, 31(3), 1999.
- [33] F Jeanpierre and F Roullier. Caractéristiques des signaux sismiques pour l'analyse des structures. Conférence CEA-EDF sur les vibrations des structures dans le domaine industriel, octobre 1979.
- [34] P.C. Jennings, G.W. Housner, and C Tsai. Simulated earthquakes motions for design purpose. volume A-1, 1969. Proc. 4th World Conference Earth. Eng. Santiago.
- [35] C Kafali and M Grigoriu. Seismic fragility analysis : application to simple linear and nonlinear systems. *Earthquake Engineering & Structural Dynamics*, 36 :1885–1900, 2007.
- [36] K Kanai. Semi-empirical formula for the seismic characteristics of the ground motion. *Bulletin of the Earthquake Research Institute, University of Tokyo*, 35(2) :308–325, 1957.

- [37] K Karhunen. Über lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.*, 37 :1–79, 1947.
- [38] K R Karim and F Yamazaki. A simplified method of constructing fragility curves for highway bridges. *Earthquake Engineering & Structural Dynamics*, 32 :1603–1626, 2003.
- [39] RP Kennedy, CA Cornell, RD Campbell, S Kaplan, and HF Perla. Probabilistic seismic safety study of an existing nuclear power plant. *Nuclear Engineering and Design*, 59 :315–338, 1980.
- [40] P.S. Koutsourelakis. Assessing structural vulnerability against earthquakes using multidimensional fragility curves : a bayesian approach. *Prob. Eng. Mech.*, 25 :49–60, 2010.
- [41] S-P Lai. Statistical characterization of strong ground motions using power spectral density fonction. *Bull. Seism. Soc. Am.*, 72 :259–274, 1982.
- [42] Y Lin and Y Yong. Evolutionary Kanai-Tajimi earthquake models. *ASCE J Eng Mech*, 113(8) :1119–1137, 1987.
- [43] Y.K. Lin. On random pulse train and its evolutionary spectral representation. *Probabilistic engineering mechanics*, 1 :219–223, 1986.
- [44] Y.K. Lin and G.Q. Cai. *Probabilistic Structural Dynamics*. McGraw-Hill, 1995.
- [45] M Loève. *Probability theory*, volume II. Graduate Texts in Mathematics. 46. Springer-Verlag, 4 edition, 1978. ISBN 0-387-90262-7.
- [46] G-C Marano, R Greco, and E Morrone. Analytical evaluation of essential facilities, fragility curves by using a stochastic approach. *Engineering Structures*, 2010.
- [47] J Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philos. Trans. Roy. Soc. London*, 83(559) :69–70, 1909.
- [48] N Metropolis, A.W Rosenbluth, M.N Rosenbluth, A.H. Teller, and E Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6) :1087–1092, 1953.
- [49] K Papadimitriou. *Stochastic characterization of strong ground motion and application to structural response*. Earthquake Engineering Research Laboratory, California Institute of Technology, Pasadena, CA, 1990. Report no. EERL 90-03.
- [50] G Perrin. *Etude du Couvert Forestier par Processus Ponctuels Marqués*. Mathématiques appliquées, Ecole Centrale Paris, 2006.
- [51] V Piluso, G Rizzano, and I Tolone. Seismic reliability assessment of a two-story steel-concrete composite frame designed according to eurocode 8. *Structural Safety*, 31 :383–395, 2009.
- [52] F Poirion and I Zentner. Non-gaussian non-stationary models for natural hazard modelling. *Applied Mathematical Modelling*, 37(8) :5938–5950, 2013.
- [53] F Poirion and I Zentner. Stochastic model construction of natural hazards given experimental data. Actes de la conférence ICOSSAR, 2013. New-york.
- [54] F Poirion and I Zentner. Stochastic model construction of observed random phenomena. *Probabilistic Engineering Mechanics*, 36 :63–71, 2014.

- [55] A Preumont. A method for the generation of artificial earthquake accelerograms. *Nuclear Engineering and Design*, 59 :357–368, 1980.
- [56] A Preumont. *Vibrations aléatoires et analyse spectrale*. Presses polytechniques et universitaires romandes, 1990.
- [57] M Priestley. Non-linear and non-stationary time series analysis. *Academic Press, Harcourt Brace Jovanovich Publishers, Great Britain*, 1988.
- [58] M.B. Priestley. Evolutionary spectra and nonstationary processes. *J. Royal Statistical Society*, 27(2) :204–237, 1964.
- [59] H Queffelec. *Topologie*. Dunod, 1998.
- [60] S Rezaeian and A Der Kiureghian. A stochastic ground motion model with separable temporal and spectral nonstationarity. *Earthquake Engineering & Structural Dynamics*, 37 :1565–1584, 2008.
- [61] S Rezaeian and A Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10) :1155–1180, 2010.
- [62] S Rezaeian and A Der Kiureghian. *Stochastic modeling and simulation ground motions for performance-based earthquake engineering*. PEER report, 2010/02.
- [63] W Rudin. *Analyse réelle et complexe*. Masson, 1975.
- [64] P Ruiz and J Penzien. Stochastic seismic response of structures. *Journal of engineering mechanics*, 94 :441–456, 1971. (ASCE).
- [65] F Sabetta and A Pugliese. Simulation of nonstationary time histories scaled for magnitudes, distance and soil conditions. 1995. 10th European Conference on Earthquake Engineering - Balkema.
- [66] F Sabetta and A Pugliese. Estimation of response spectra and simulation of nonstationary earthquake ground motions. *Bulletin of Seismological Society of America*, 86(2) :337–352, 1996.
- [67] N Saha and A Naess. Monte carlo-based method for predicting extreme value statistics of uncertain structures. *Journal of Engineering Mechanics ASCE*, pages 1491–1501, Dec. 2010.
- [68] G Saporta. *Probabilités, analyse des données et statistique*. Editions Technip, 2006. 2e édition.
- [69] M Shinozuka. Stochastic fields and their digital simulations. *Stochastic methods in structural dynamics*, pages 93–133, 1987. Martinus Nijhoff Publishers, Dordrecht.
- [70] M Shinozuka, M Q Feng, H Kim, T Uzawa, and T Ueda. *Statistical Analysis of Fragility Curves*. Department of Civil and Environmental Engineering - University of Southern California, 2001. Technical Report MCEER.
- [71] M Shinozuka, M.Q. Feng, H.K. Kim, and S.H. Kim. Nonlinear static procedure for fragility curve development. *ASCE Journal of Engineering Mechanics*, 126(12) :1267–1295, 2000.
- [72] BW Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.
- [73] C Soize. *Méthodes mathématiques en analyse du signal*. Masson, 1993.

- [74] C Soize. Identification of high dimension polynomial chaos expansions with random coefficients for non-gaussian tensor-valued random fields using partial and limited experimental data. *Computer methods in applied mechanics and engineering*, 199 :2150–2164, 2010.
- [75] C Soize and R Ghanem. Physical systems with random uncertainties : chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing*, 26(2) :395–410, 2004.
- [76] P.D. Spanos and L.M. Vargas Loli. A statistical approach to generation of design spectrum compatible earthquake time histories. *Soil dynamics and earth. eng.*, 4(1) :2–8, 1985.
- [77] D Straub and A Der Kiureghian. Improved seismic fragility modelling from empirical data. *Structural safety*, 30 :320–336, 2008.
- [78] H Tajimi. A statistical method of determining the maximum response of a building structure during an earthquake. volume 2, pages 781–798. 2nde WCEE, Tokyo, 1960.
- [79] E.H. Vanmarcke. Properties of spectral moments with applications to random vibration. *J. Engrg. Mech. Div. ASCE*, 98 :425–446, 1972.
- [80] EH Vanmarcke. On the distribution of the first-passage time for stationary random processes. *Journal of applied mechanics*, pages 215–220, 1975.
- [81] E.H. Vanmarcke and D.A. Gasparini. Simulated earthquake ground motions. 1977. Proc. 4th Int. Conf. on Smirt, K1/9, San Francisco.
- [82] N Wiener. The homogeneous chaos. *Am. J. Math.*, 60 :897–936, 1938.
- [83] C.H. Yeh and Y.K. Wen. Modeling of non-stationary ground motion and analysis of inelastic structural response. *Structural safety*, 8(1-4) :281–298, 1990.
- [84] I Zentner. Modélisation de la turbulence atmosphérique. application aux éoliennes. Dec. 2002.
- [85] I Zentner. Méthodes probabilistes dans les analyses sismiques : modélisation, propagation et hiérarchisation des incertitudes. EDF R&D, 2007. 7e Colloque National AFPS - Ecole Centrale Paris.
- [86] I Zentner. Numerical computation of fragility curves for npp equipment. *Nuclear Eng. Design*, 240(6) :1614–1621, 2010.
- [87] I Zentner. Génération de signaux sismiques. Document interne EDF Code Aster, 2014.
- [88] I Zentner, L D’Amico, and P Cacciola. Simulation of non stationary ground motion compatible with nga-spectra. Actes de la conférence ICOSAR, 2013. New-york.
- [89] I Zentner and F Poirion. Enrichment of seismic ground motion databasis using karhunen-loève expansion. *Earthquake Engineering & Structural Dynamics*, 41(14) :1945–1957, 2012.

Notations

Dans l'intégralité de ce mémoire, il est fait usage :

- des minuscules pour les scalaires, réels et complexes ; des minuscules en **gras** pour les vecteurs ;
- des majuscules pour les matrices et les variables aléatoires ; des majuscules en **gras** pour les vecteurs aléatoires.

\mathcal{B} tribu borélienne de \mathbb{R}

\mathcal{B}^N tribu borélienne de \mathbb{R}^N

\mathbb{C} corps des complexes

dt (ou dx) mesure de Lebesgue sur \mathbb{R}

e constante de Néper

$E[X]$ espérance d'une variable aléatoire X

$g \simeq 9.81 \text{ m.s}^{-2}$ l'accélération de la pesanteur

I_N matrice unité de $\mathcal{M}_N(\mathbb{R})$

\mathbb{I}_A fonction indicatrice de l'ensemble A

$L_{\mathbb{R}}^p(X, \mathcal{T}, \mu)$ espace vectoriel des (classes modulo les fonctions nulles dt -presque partout de) fonctions réelles de l'espace mesuré (X, \mathcal{T}, μ) p -intégrables au sens de Lebesgue

$L_{\mathbb{R}}^p(A, dt)$ espace vectoriel des (classes modulo les fonctions nulles dt -presque partout de) fonctions réelles définies sur une partie A de \mathbb{R} et p -intégrables au sens de Lebesgue

$\mathcal{LN}(A_m, \beta)$ loi log-normale de médiane A_m et d'écart-type logarithmique β

$\mathcal{M}_{n,p}(\mathbb{K})$ espace vectoriel des matrices de taille $n \times p$ à coefficients dans le corps \mathbb{K}

M^T transposée de la matrice M

$[M]_{ij}$ coefficient i, j de la matrice M

\mathbb{N} ensemble des entiers naturels

$\mathcal{N}(\mu, \sigma^2)$ loi normale d'espérance μ et de variance σ^2

$\mathcal{N}(\mu, \Gamma)$ loi multinormale d'espérance μ et de variance-covariance Γ

$O_n(\mathbb{R})$ groupe des matrices orthogonales réelles de taille n

\mathbb{R} corps des réels	$\mathcal{U}(I)$ loi uniforme continue sur l'intervalle I
$\Re[z]$ partie réelle d'un nombre complexe z	v.a. variable aléatoire
$\mathcal{S}_n(\mathbb{R})$ espace vectoriel des matrices symétriques réelles	$V(X)$ variance d'une variable aléatoire X
$\mathcal{S}_n^+(\mathbb{R})$ espace vectoriel des matrices symétriques réelles positives	z^* conjugué du nombre complexe z
$\mathcal{S}_n^{+\star}(\mathbb{R})$ espace vectoriel des matrices symétriques réelles définies positives	π constante d'Archimède
$\mathcal{S}_{\mathbb{K}}(A)$ spectre d'une matrice A sur le corps \mathbb{K}	$\rho(A)$ rayon spectral d'une matrice A
$\mathcal{U}(a,b)$ loi uniforme discrète sur le segment $\llbracket a; b \rrbracket$	σ_X écart-type d'une variable aléatoire X
	$\chi_A(X)$ polynôme caractéristique de A
	§ paragraphe

